UNIVERSIDAD AUTÓNOMA DE MADRID

STUDY OF THE ENHANCEMENT OF COMPUTATIONAL CAPABILITIES IN AN OLFACTORY SYSTEM MODEL BY MEANS OF NEURONAL DIVERSITY AND GAIN CONTROL MECHANISM

Thesis presented by Aarón Montero Montero
for the degree of Ph.D. in Computer Science
Supervised by Dr. Francisco de Borja Rodríguez Ortiz

Madrid 2019

Dedicated to
Manuel Montero Toledo
(1946-2017)

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Prof. Francisco de Borja Rodríguez for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. I hope that this thesis, fruit of years of research, can be worthy of all the time, dedication and trust that he has placed in me.

My sincere gratitude also goes to Dr. Ramón Huerta for offering me the internship opportunities in his group and leading me working on diverse exciting projects. Without his precious support, it would not be possible to conduct this research.

Besides, I would like to thank my thesis committee members: Prof. Pablo Varona, Prof. Jorge Mejías, and Prof. Mario Salvador González for offering to evaluate this work, for their encouragement, insightful comments, and hard questions, which incented me to widen my research from various perspectives.

I thank my labmates in Grupo de Neurocomputación Biológica and especially those with whom I have shared most moments in these years: Jesús Díaz, Fernando Herrero, Miguel Urizar, Carlos Saura, Alejandro Pequeño, Angel Lareo, Irene Elices, Victor H. García, Vinicio Changoluisa, Guillermo Sarasa, Angel Fuente, Manuel Reyes, Rodrigo Amaducci and Jessica López-Hazas. Thank you for your support and for all the good times we have shared. You are the best!

I also want to thank my friends in Barcelona, especially those who shared the research path and we have been able to support each other: Jose Luis Villanueva, Elena Sanchez, Ruth Ayllon, Borja Ordoñez.

Last but not least, I would like to thank my parents Manuel Montero and Josefa Montero. I cannot imagine better parents than you and I thank the fortune of being your son and receiving from you all the love and care. You are the reason behind this thesis due to how much you have taught me the value of knowledge. Mom, I know how much it hurt you not to be able to study when you were a girl and how much insecurity you had because of it. I know how much effort you put in instilling to my sister and me the importance of studying and how many sacrifices you made

so that we could go to the university. You always wanted that we could learn from dad and take him as a model, but you cannot imagine how much you taught through your commitment to give us everything you could not have. Dad, you would not believe how many times I have told the story of your life, how a country boy from a poor family could get three academic degrees, and I have boasted of you because of it. You always told me that I was smarter than you, but I always envied and admired your motivation, perfectionism and hard work. All my achievements seem so easy compared to yours that I wonder if someday I could make the best of myself, as you did. In fact, if someone asked me what I want to be in the future, I would say that I want to be like you. I hope to be a father as great as you and to awaken in my children the same admiration that I have for you. For all of this, I want to dedicate to both of you this thesis and share with you my Ph.D title. I love you so much!

Contents

Abstract

Every time you inhale, you introduce molecules from the air inside your body. Some of these molecules interact with neuronal receptors located in the olfactory epithelium, which generates a nerve impulse that travels through the olfactory tract to your brain. A direct trip unlike that carried out by the rest of the senses, that have to go through the thalamus in the first place. Hence, odor information can be perceived earlier, alerting us, for example, of potential dangers such as something burning close to us. This fast processing together with the human capability of recognizing a trillion odorants, makes the olfactory system a tool of great interest to study, since it could bring great innovations to the field of pattern recognition and machine learning.

However, instead of simulating computationally the human olfactory system, we have focused on a simpler and better known system from the point of view of biology, the insect olfactory system. This system allows us to study what properties of its neural network are linked to its ability to classify odorants and extract conclusions that can be extrapolated to our own olfactory system.

The objective of this thesis is to analyze the role of neuronal heterogeneity in odor discrimination through the computational model of the olfactory system of insects. Specifically, we will analyze the role of three types of neuronal heterogeneity observed in this olfactory system, which are related to each other. In the first place, the existence of varying neural thresholds, which we have demonstrated that allow improving the classification results with respect to the use of the same neural threshold for all neurons. Secondly, specialist and generalist neurons, which respond differently to input stimuli because of their neural threshold values. For this neuronal diversity we observed that although specialist neurons are crucial for the classification of patterns, there is an optimal ratio of specialists/generalists for this classification based on the complexity of the input patterns. Finally, the gain control mechanism of the antennal lobe of insects, produced by the interaction of two populations of inhibitory neurons. We observed that using this inhibitory heterogeneity, to model this gain control mechanism, we are able to simulate more properly the biological behavior of this mechanism and obtain a better classification result with respect to modeling it through a homogeneous inhibition.

Thus, because whenever we apply the neural heterogeneity in our computational model we obtain better classification results, we suggest that including these mechanisms in artificial neural networks, chemical sensors and other tools for solving pattern recognition problems can be beneficial.

Resumen

Cada vez que inhalas atraes hacia tu interior moléculas que se encuentran en el aire. Algunas de estas moléculas al hacer contacto con los receptores neuronales situados en el epitelio olfatorio desencadenan un impulso nervioso que viajara a través del tracto olfatorio hasta tu cerebro. Un viaje directo a diferencia del realizado por el resto de sentidos, los cuales han de pasar previamente por el tálamo. Esto hace que su información sea percibida antes, alertándonos, por ejemplo, de ciertos peligros cómo puede ser de algo quemándose. Este rápido procesamiento unido a que se estima que el ser humano es capaz de reconocer un billón de olores, hace del sistema olfativo una herramienta de gran interés para su estudio, puesto que podría traer gran innovación al campo del reconocimiento de patrones y el aprendizaje automático.

No obstante, en lugar de simular computacionalmente el sistema olfativo humano, nos hemos centrado en un sistema más simple y mejor conocido desde el punto de vista de la biología, el sistema olfativo de los insectos. Este sistema nos permite estudiar que propiedades de su red neuronal que están vinculadas con su capacidad de clasificación y extraer conclusiones extrapolables a nuestro propio sistema olfativo.

El objetivo de esta tesis es analizar el papel de la heterogeneidad neuronal en la discriminación de odorantes a través del modelo computacional del sistema olfativo de los insectos. Concretamente, analizaremos el papel de tres tipos de heterogeneidad neuronal observadas en dicho sistema olfativo las cuales se hallan relacionadas entre sí. En primer lugar, la existencia de umbrales neuronales heterogéneos, los cuales hemos demostrados que permiten mejorar los resultados de clasificación respecto al uso de un mismo umbral neuronal para todas las neuronas. En segundo lugar, las neuronas especialistas y generalistas, las cuales responden de forma heterogénea a los estímulos de entrada a causa de sus valores de umbral neuronal. Mostrándose que si bien las neuronas especialistas son vitales para la clasificación de los patrones, existe una proporción óptima de especialistas/generalistas para dicha clasificación en función de la complejidad de los patrones de entrada. Finalmente, el mecanismo de control de ganancia presente en el lóbulo

antenal de los insectos, producido por la interacción de dos poblaciones de neuronas inhibitorias. Observando que si utilizamos esta heterogenidad inhibitoria, en el modelo del mecanismo de control de ganancia, logramos simular más fielmente el funcionamiento biológico de este mecanismo y obtenemos un mejor resultado de clasificación respecto a modelarlo a través de una inhibición homogénea.

Por tanto, debido a que siempre que aplicamos la heterogeneidad neuronal en nuestro modelo computacional obtenemos mejores resultados de clasificación, sugerimos que incluir estos mecanismos en redes neuronales artificiales, sensores químicos y otras herramientas para resolver problemas de reconocimiento de patrones puede ser beneficioso.

CHAPTER

1

Introduction

Smell is the most primitive sense because of its direct relationship with the essential needs for animal life, such as nutrition and reproduction. Perhaps, due to being the first sense developed, the olfactory pathways appear in the embryo before the thalamus [194]. This entails that olfactory stimuli are the only sensory stimuli that go directly to the brain without going through this relay center [64], accelerating their processing. A complex treatment, because odors are generally blends of several chemical components, whose perception can be hindered by changes in their ratios, the background odor context and the intermittency with which they are received due to variations in the fluid dynamics that transports them (air or water) [121]. However, despite these difficulties, the olfactory system is not only able to generalize the information it receives to identify an odor regardless of environmental conditions, but it is able to identify a large number of different odors. According to recent studies, humans are able to recognize a trillion odorants (10^{12}) [18], even though we do not have one of the most developed olfactory systems of the animal kingdom (humans have 396 olfactory receptor genes compared to the $1,948$ genes of number one, the African elephant [158]). This potential for pattern recognition tasks is what has led us to study it. Nevertheless, simulating computationally the human olfactory system and studying what properties of its neural network are linked to its ability to classify odorants is beyond our possibilities.

Thus, we have focused on the olfactory system of insects (see panel (b) in Fig. 1.1), which has a simple structural organization [5, 43, 58, 59, 78, 95, 140, 209] that is easier to model computationally and whose structural similarities with the rest of animals will allow us to extrapolate the conclusions reached in this work.

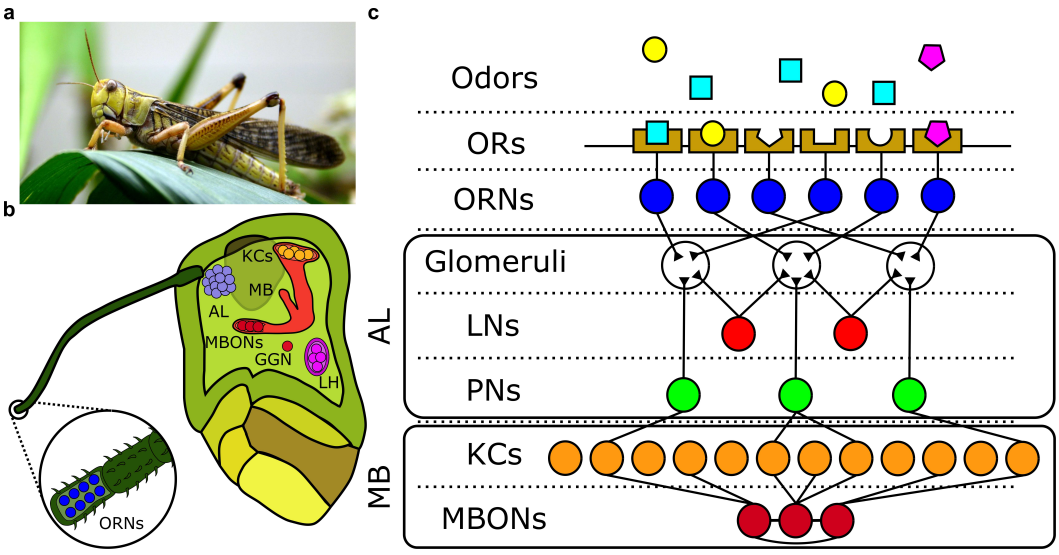


Figure 1.1: Structure of the insect olfactory system. Panel (a) shows a locust (picture extracted from [232]). Panel (b) shows a representation of its olfactory system. Panel (c) shows a schematic of this system with the different parts involved in the processing and recognition of odor patterns. In the antenna, ORNs detect the odorants and send the odor stimuli to AL for its encoding. ORNs that are sensitive to a certain odor send the stimuli to a single glomerulus of AL, which transmits it to their PNs and LNs. Of these neurons, only PNs transmit the odor stimuli to deeper areas of the brain such as the MB, which is responsible for the learning and recognition of odor patterns. Inside MB, the stimuli are received by its KCs, which increase the separability of odor stimuli and facilitates the final task of odor classification in MBONs.

The odor processing in the insect olfactory system (see panel (c) in Fig. 1.1) begins at the antennae, where a massive number of olfactory receptor neurons (ORNs) send the odor stimuli to the antennal lobe (AL) for their encoding. The AL is composed by densely packed neuropils [94], called glomeruli. Optical imaging of intracellular calcium concentration in insects suggests that ORNs responsible for detecting a certain odorant are likely to converge generally onto a single glomerulus [20, 44, 55, 57, 60, 203, 229]. Hence, glomeruli serve as collectors of a specific

odor stimulus, which is transmitted to the two types of AL neurons that innervate them, the projection neurons (PNs) and the local neurons (LNs).

However, the transmission of odor information suffers, in this part of the olfactory system, an alteration. While odor concentration affects the activity of ORNs, the signal transmitted by the PNs is invariant to it [84, 207]. This is due to the gain control mechanism [191, 201] realized by LNs that controls the activity of PNs by inhibiting them [167, 168, 235]. The inhibition produced by LNs seems to be heterogeneous due to the existence of two different types of neurons: homoLNs, which innervate most if not all glomeruli uniformly, and heteroLNs, which innervate only a few of them [65, 193]. This may entail that the gain control observed in AL is not only due to the fact that there is inhibition by LNs, but that this inhibition is heterogeneous, as suggested in [137]. Once this gain control has been carried out, the PNs transmit the odor stimuli to deeper layers of the insect brain, varying its destination according to the type of PN: uniglomerular or multiglomerular. PNs are usually uniglomerular (which have dendrites in a single glomerulus) in most species. For example, the proportion of multiglomerular PNs with respect to uniglomerular PNs is below 10% in *Drosophila* (fruit fly). Uniglomerular PNs innervate the mushroom body (MB) and the lateral horn (LH) [198, 200, 207]. On the other hand, multiglomerular PNs (which branch in several, if not all, glomeruli) generally do not innervate the MBs, transferring the odor information to other areas of protocerebrum instead. Between these two types of PNs we will focus on the uniglomerular ones, since they are the ones that transmit the odor information to the MB, which is involved in the learning and recognition of the odorants.

The stimuli sent by PNs to the MB is received by the Kenyon cells (KCs). The neural connections between PNs and KCs are random (with a connection probability between 0.1 and 0.5 [37, 62, 101, 192]) since its connectivity pattern does not show reproducibility between individuals of the same species [21, 132, 237]. This randomness, together with the connection divergence from PNs to KCs (with an expansion ratio of 1:10 in the *Drosophila* [153, 223] and 1:50 in the locust [164, 175, 243]), suggests that the function of KCs is to increase the separability of odor stimuli. At this point, we can consider that the spatio-temporal odor patterns from PNs [109, 159, 164] are converted into spatial-only patterns in KCs because of

their activity pattern. KCs are inactive most of the time [175, 213], with a mean firing frequency lower than 1 Hz, but when they finally fire, their neuronal response is produced by the coincidence of concurrent spikes from the PNs followed by a reset [107, 231, 240]. This activity pattern has two direct consequences. The first, and most obvious, is that the odor coding becomes sparse, which facilitates the separability and subsequent classification of odor patterns [101, 117]. The second is that, given the inactivity periods of KCs before and after their response to stimuli, KCs take a spatial "snapshot" of the odor pattern, which is all the MB needs to recognize the odorant [110, 206]. To maintain this low activity and sparse coding, KCs seem to use two mechanisms. On the one hand, the inhibition received by GABAergic neurons [12, 131, 172, 242], which use the activity received from KCs to inhibit them. On the other, the heterogeneity in their firing thresholds [2, 174, 175], that would help them adapt more quickly to the stimuli they receive and would be responsible for the great diversity observed in their neural sensitivity [182] (which allows differentiating between specialist and generalist neurons [28, 102]).

Finally, the odor information reaches the MB output neurons (MBONs) through connections established by synaptic plasticity and that storage the olfactory associative memories [22, 23]. These MBONs are subjected to local inhibition [199] that may underlie competition among them and the winner selection of a multiclass problem [169].

1.1 Objectives

The insect olfactory system presents different kinds of neural heterogeneity. Although neural heterogeneity is not usually applied in biological models, their existence in biological systems has been reported as beneficial for information processing. Experiments in mammals [81, 170] and electric fishes [7, 130, 195] showed that neural heterogeneity improves the coding efficiency. For example, the heterogeneous response properties of inferior colliculus neurons in mice contribute to the efficient coding of behaviorally relevant vocalizations [81]. Thus, the objective of this thesis is to analyze, in an insect olfactory model, the role of neural heterogeneity in the processing of odor patterns and its effects on their clas-

sification. In particular, we focus on analyzing the three neural heterogeneities mentioned during the description of the insect olfactory system:

- **The variability of firing thresholds in KCs.** We analyze the utility of using heterogeneous thresholds to solve pattern recognition problems. This analysis is carried out by comparing the classification results of different types of patterns when, in our computational model, we use the same threshold value for all neurons (homogeneous thresholds) or different threshold values (heterogeneous thresholds).

- **The diversity observed in the neural sensitivity of KCs because of their heterogeneous thresholds.** In terms of neural sensitivity, two kinds of neurons are always differentiated based on their responses to stimuli: specialists and generalists. Specialist neurons are selective responding to stimuli, while generalists code for multiple stimuli [28]. The role of both classes of neurons in the olfactory system is still unclear [28, 102]. However, it is suggested that specialists are crucial for discrimination, while generalists play a key role in extracting and discovering common features [236]. We hypothesize that this heterogeneity in neural responses is an animal adaptation to the complexity of stimuli in the natural environment. Thus, our aim is to analyze the ratio of specialists and generalists required by the neural system as a function of the stimulus complexity.

- **The inhibitory heterogeneity that produces the gain control mechanism observed in AL.** Our goal is to progressively develop a gain control mechanism that can be plausible from a biological point of view. For this purpose, we base on its biological structure and the input-output activity relationship observed in its PNs.

1.2 Outline

The thesis is structured as follows:

- **Chapter 2** introduce the related work, which explains some important concepts and definitions that will be used afterwards: the nature of the odor

stimulus, the different parts of the insect olfactory system involved in its processing and how this system can be represented by computational modeling

- **Chapter 3** presents the materials and methodology used during the development of the thesis. First, we describe the characteristics of the computational model used, such as the neural network and neuron models and the learning algorithms. Second, the design of experiments to study the firing thresholds, the populations of specialist and generalist neurons and the gain control mechanism observed in AL. Finally, we show the patterns used in these experiments.

- **Chapter 4** collects the results obtained on the role of the three neuronal heterogeneities studied during odor processing.

- **Chapter 5** and **Chapter 6** present the discussions and conclusions reached for the different objectives stated at the beginning of this thesis in English and Spanish.

- **Chapter 7** lays out the different studies that emerged from this thesis that will be accomplished in the near future.

- **Chapter 8** and **Chapter 9** show the dissemination of the research carried out during the thesis through publications and international congresses.

- **Chapter 10** describes the research internships accomplished.

- **Appendix A** contains a glossary of the acronyms used.

- **Appendix B** describes the algorithms of threshold selection used throughout the thesis.

- **Appendix C** details the measure used to calculate the complexity degree of input patterns.

CHAPTER

2

Related work

Since the aim of this thesis is to study the properties of the olfactory system of insects that can be applied to computer science to improve learning and pattern recognition techniques, in this chapter we will take a tour through the different areas of knowledge involved.

In section 1 we focus on the nature of the stimulus itself, how it is detected by the olfactory system and how the coding of odor patterns changes as it is transmitted throughout this system. This section serves to argue decisions on the computational model, such as the choice of the McCulloch-Pitts neuron model to simulate the hidden and output neurons (see section 3.1.2) that represent the mushroom body (MB), and the use of Gaussian distributions (see section 3.3.2) to emulate the activity in the glomeruli of the antennal lobe (AL), which represent the input layer of our model.

Section 2 describes the insect olfactory system. This information is used for developing the computational model (see section 3.1), which is focused on those parts related to the odor classification. Among the characteristics that we took to make our computational model are the dimensions of the different layers of the neural network, the connectivity patterns between layers, the gain control mechanisms and the lateral inhibitions between neural populations.

Section 3 mentions different computational models based on the olfactory system, focusing on those that represent more widely the olfactory system of insects and reasoning the use of the computational model detailed in [89, 90] as the basis for the one used in our study (see section 3.1).

Finally, section 4 is focused on neural heterogeneity. We provide some reasons why this property observed in the insect olfactory system deserves to be studied and the knowledge we have about the three types of heterogeneity analyzed in this thesis: the variability of neural thresholds, the variability of response to stimuli of the KCs and the heterogeneous inhibition of the LNs (see section 3.2).

2.1 Olfactory stimulus

Sensory receptors provide information about internal and external events. This information allows living beings to adapt to their environment and survive in it. The receptors are commonly characterized in four distinct categories: chemoreceptors, photoreceptors, mechanoreceptors, and thermoreceptors. Of all of them, chemoreceptors, in the form of olfactory receptors (ORs), are the most primitive because they are strongly tied to the search of food, threats, and mates. This vital information is transmitted by its source through odor plumes [226] (Fig.2.1) whose variable structure can sometimes be visually observed as in the case of smoke. The odor plume structure is affected directly by the fluid properties through which they are propagated. These change the distribution of odorous molecules in time and

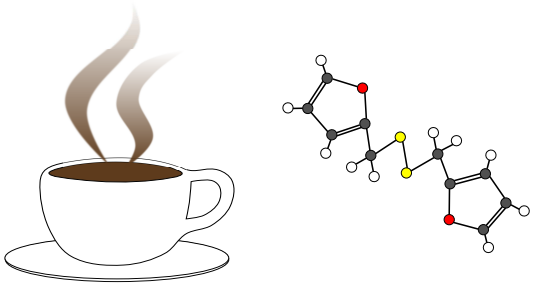


Figure 2.1: Odor representation. Odorants spread from their source through odor plumes. Each odor has its own agonist (a chemical compound that activates the OR) [33]. In the case of coffee, the agonist is coffee difuran ($C_{10}H_{10}O_2S_2$).

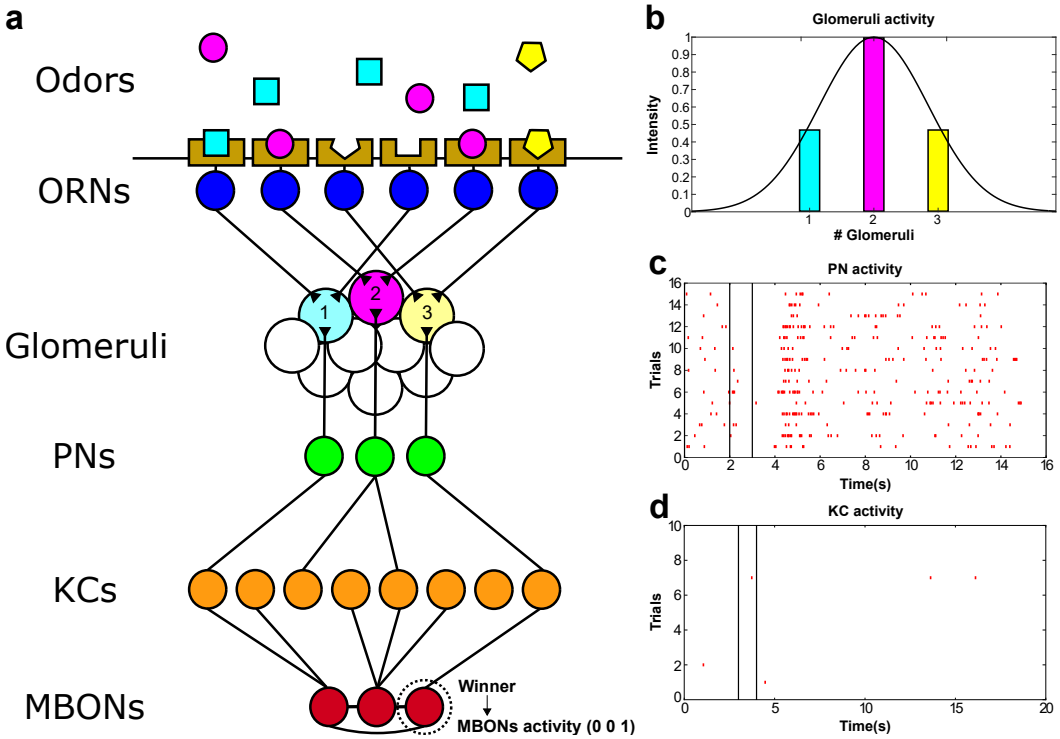


Figure 2.2: Odor coding through the insect olfactory system. Panel (a) shows the different areas of the insect olfactory system involved in the odor pattern recognition process. Panel (b) shows a representation of glomeruli activity as a function of the olfactory stimulus shown in panel (a). Panels (c) and (d) show the activity observed in a PN and a KC based on Perez-Orive recordings [175] using the toolbox developed by Jessica Lopez-Hazas [122]. In these panels, we observe that while the PNs shows a high activity in the different trials, the KCs is mostly silent and transforms the odor pattern in sparse code.

space, which implies the reception of a highly intermittent signal when the odor is patchily distributed.

When we inhale, we capture odorous molecules transported by these plumes. The chemical compound or compounds that compose each odor have two ways to bind to ORs: as agonists or antagonists [165]. A chemical compound that binds to an OR as an agonist will activate it, generating a nerve impulse. On the other hand, an antagonist that binds to a receptor will not activate it but block it, preventing agonists from activating it. However, regardless of knowing that certain chemicals activate or block ORs after binding to them, how ORs actually work remains un-

known. There are two theories: docking and vibration. The docking theory [142], also known as shape theory, considers that ORs have a specific shape that fits with the chemicals that recognizes, what is known as ”lock and key” model. The problem with this theory is that it is not possible to have one type of OR for each odorant that we recognize (remember that with 396 different types of ORs, we recognize a trillion odorants). To solve this problem, the odotope theory [151], or weak shape theory, was proposed. This recent theory suggests that each receptor fits with one of the sections of the odor molecule, being the brain responsible for processing the combined signal into an interpreted smell. This would explain why chemical compounds that have similar sections (for example, a sulfur hydrogen bond) have a similar odor (rotten eggs, for the given example). However, this theory also has its problems, since there are molecules composed of the same molecular groups but arranged in different ways that produce a different smell, such as vanillin and isovanillin. Since these compounds can be differentiated through their resonance frequency, the vibration theory [222] was proposed. This controversial theory is based in inelastic electron tunneling, by which ORs are waiting for an odor molecule with a certain resonance frequency that allows the electron to pass through the gap. To support this theory, odor molecules were modified by replacing the hydrogen molecules with deuterium to vary their resonance frequencies without varying their chemical properties. Experiments in humans [61] and in *Drosophila* [50] showed that the original and modified molecules had different odors. However, attempts to replicate these results by other researchers found no differentiation in the smell [13, 104]. Because of this, the docking theory is considered the most plausible [189].

It is at this point, when the olfactory stimulus becomes a series of nerve impulses that encode the odorant, where our interest really begins.

As we mentioned in the previous chapter, ORNs with the same type of OR transmit their information to the same glomerulus of AL (see panel (a) in Fig.2.2). This implies that different olfactory stimuli (composed by one or multiple chemical compounds) activate different glomeruli [56] . The concentration level of an odorant can affect the number of glomeruli that respond to it, as well as the intensity of their response (see Fig.1a in [208] and in [53]). Hence, we can consider odor coding in the glomeruli as a Gaussian distribution [103, 192], in which the

activity is focused on those glomeruli more closely related to the odorant and the concentration level determines the activity intensity and the total number of glomeruli stimulated (see panel (b) in Fig.2.2). Although the odor coding in PNs has spatio-temporal dimensions (panel (c) of Fig.2.2) and different odorants produce different temporal-response patterns [112], we used spatial patterns (see section 3.3) as representation of the olfactory stimuli in PNs, such as Gaussian distributions, because KCs seem to make "snapshots" of the odor patterns [110, 206]. This loss of the temporal dimension in KCs is due to the fact that they act as coincid-

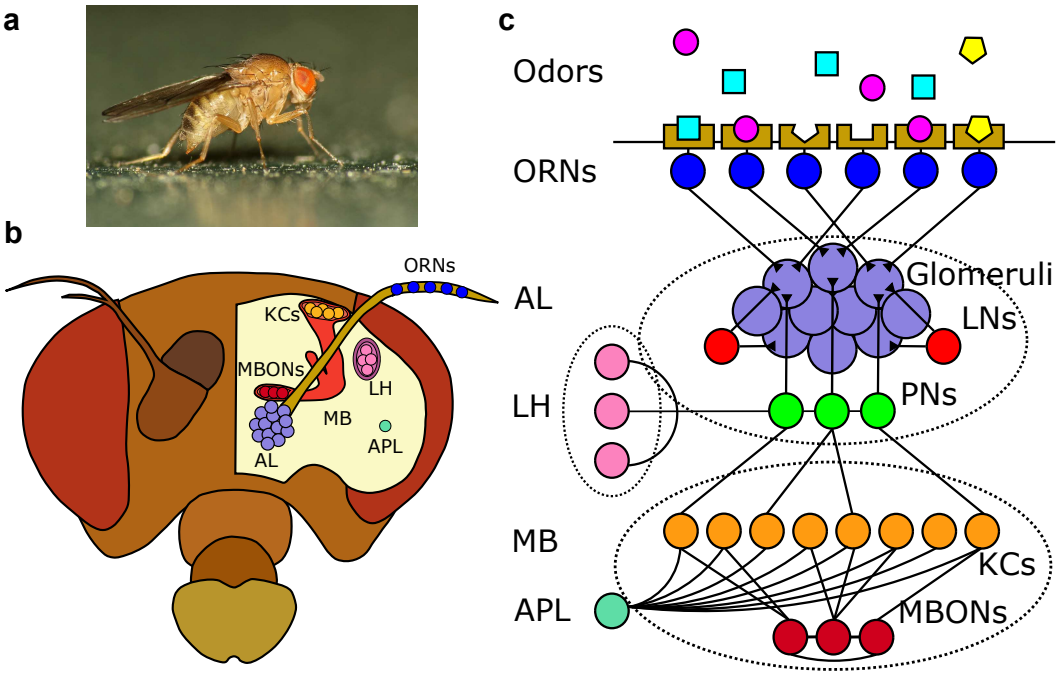


Figure 2.3: Olfactory system of insects. Panel (a) shows a *Drosophila* (picture extracted from [93]). Panel (b) displays the different parts of its olfactory system. Panel (c) presents a schematic of this system. ORNs located on antennae sense odors and send their output to LNs and PNs in the AL. Inhibitory and excitatory interactions between ORNs, LNs, and PNs help transform the sensory information into a spatio-temporal code, which is then carried by PNs to MB and LH. In MB, the information is received by a large population of sparsely responding KCs, which then converge on the MBONs. In *Drosophila*, APL helps maintain sparse responses in KCs by forming a normalizing feedback loop; other insects have similar feedback mechanisms (such as GGN in the locust).

ence detectors that are mostly in silence. This silent behavior can be observed in panel (d) of Fig.2.2 together with the sparse coding used by these neurons, which facilitates the separability of odor patterns (see Fig.2 in [206]). Once the olfactory stimulus reaches the MBONs, the odorant is finally identified by combining the synaptic plasticity of the connections between KCs and MBONs [22, 23] and the lateral inhibition [199] between the latter (which works similar to the winner-takes-all voting [169], see panel (a) in Fig.2.2).

2.2 Insect olfactory system

The olfactory system of insects (shown in Fig.2.3) is composed by four essential parts: Olfactory Receptor Neurons (ORNs), Antennal Lobe (AL), Mushroom Body (MB) and Lateral Horn (LH). Although we are going to describe these four parts below, in our computational model we only modeled two of them, namely AL and MB, those in responsible for coding and classifying the odorants.

2.2.1 Olfactory receptor neurons

The chemical compounds, which compose the odorants, are received by ORNs located within sensilla, hair-like structures that extend from the insect cuticular surface on olfactory appendages [17, 19, 27, 228] (see Fig.2.4). Odors molecules diffuse through pores in the sensilla walls, entering the sensillum lymph where they interact with odorant-binding proteins (OBPs) and are transferred through the aqueous medium towards the dendrites of ORNs [114, 227]. Adequate compounds induce a depolarization of responsive neurons leading to action potentials [154]. The transduce odorant binding to cellular excitation in the ORNs is made by the odorant receptors (ORs), which are mostly related to seven transmembrane G-protein-coupled receptors (GPCRs). When a G-protein reacts to an odorant, this activates the Adenylate Cyclase (AC), resulting in the generation of cyclic Adenosine Monophosphate (cAMP). The increase in cAMP opens ionic channels that permit Na^+ and Ca^{2+} (mostly Ca^{2+}) entry through the cellular membrane, thus depolarizing the neuron [138] (see Fig.2.5). Every ORN typically expresses only

one type of OR. ORNs that express the same type of OR converge on a single glomerulus in the antennal lobe (AL) [20, 44, 55, 57, 60, 118, 203, 229].

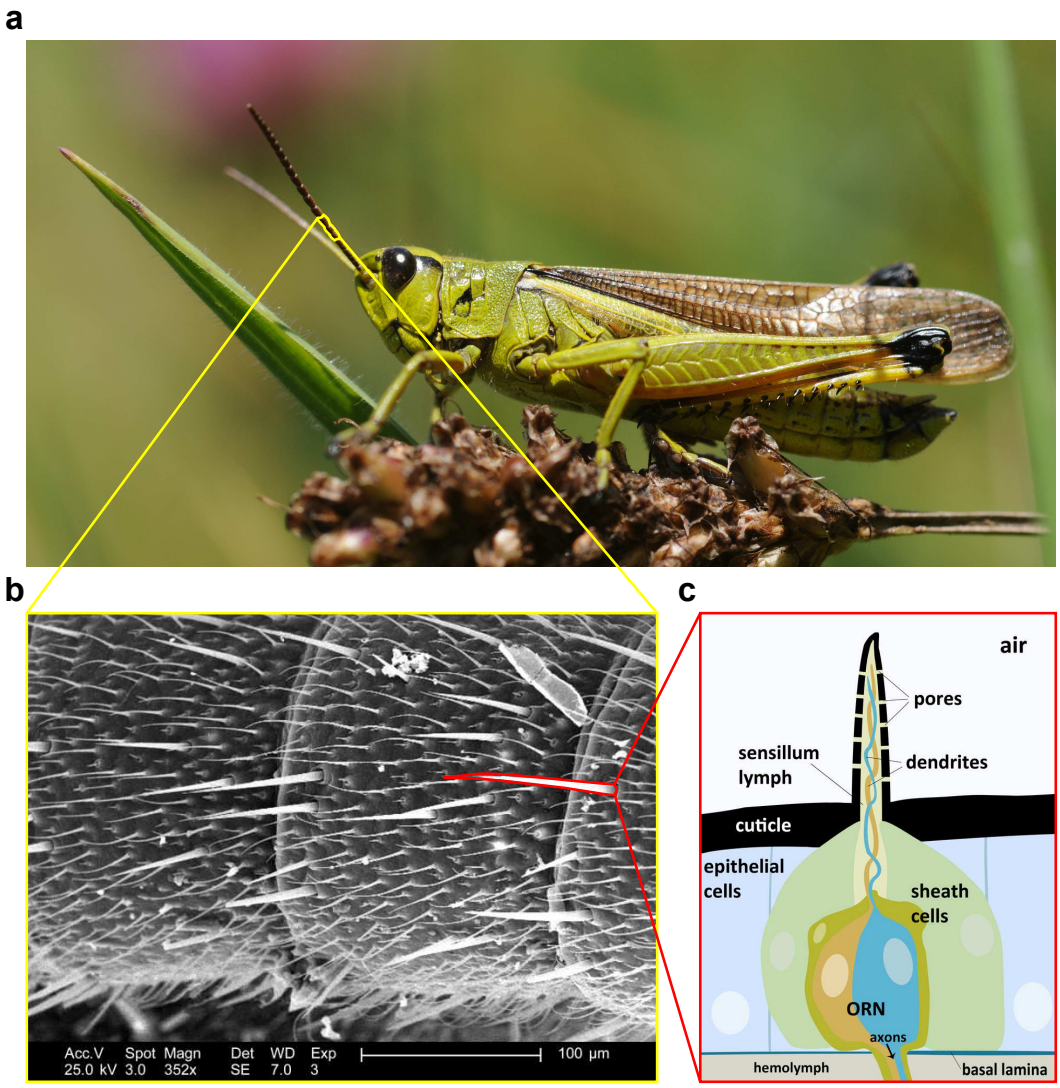


Figure 2.4: Antennae structure of insect. Panel (a) shows a locust (modified picture extracted from [15]). Panel (b) displays a scanning electron micrograph of the antenna, which displays the external morphology of sensilla (modified picture extracted from [97]). Panel (c) shows the schematic of an olfactory sensillium (modified picture extracted from [1]).

2.2.2 Antennal lobe

The antennal lobe is composed of densely packed neuropils, named glomeruli, where the ORNs synapse with projection neurons (PNs) and local neurons (LNs) [14, 63]. Optical imaging of intracellular calcium concentration in insects suggests that ORNs expressing the same OR are likely to converge generally on a single glomerulus [20, 44, 55, 57, 60, 118, 203, 229].

Regarding PNs, they have both input and output synapses within olfactory glomeruli. PNs are usually uniglomerular (having dendrites into a single glomerulus) in most species (for example, the proportion of multiglomerular PNs with respect to uniglomerular PNs is below 10% in *Drosophila*). On the one hand, uniglomerular PNs innervate both the MBs and LH [207], forming two distinct

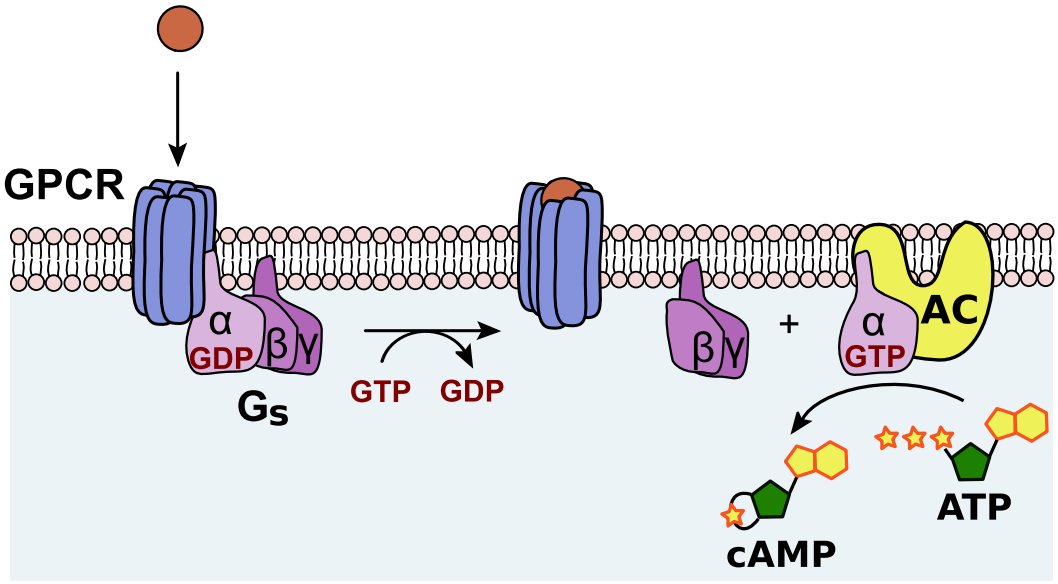


Figure 2.5: Process of electrically encoding of a chemical compound in the ORNs. Odorants bind to the G-protein-coupled receptors (GPCRs), which stimulates the α subunit. The α subunit drops Guanosine Diphosphate (GDP) and gain Guanosine Triphosphate (GTP). This activates the Adenylate Cyclase (AC), resulting in the generation of cyclic Adenosine Monophosphate (cAMP) from Adenosine Triphosphate (ATP). The increase in cAMP opens ionic channels that permit Na^+ and Ca^{2+} (mostly Ca^{2+}) entry through the cellular membrane, thus depolarizing the neuron (modified picture extracted from [155]).

antenno-cerebral tracts: one located close to the brain midline, known as middle antenno-cerebral tract (mACT) in bees and inner antenno-cerebral tract (iACT) in other species, and the other located laterally (lACT). On the other hand, multiglomerular neurons branch in several if not all glomeruli, and generally do not innervate the MBs, but a variety of other areas in the protocerebrum using a group of smaller, intermediate tracts. In the case of locust, the organization of ACTs is clearly different from that in other insects. In locusts, only a single ACT close to the brain midline connects the ALs to the MB [116], and a few minor tracts connect the AL to other areas [92]. PNs are not uniglomerular, but branch in a limited number of glomeruli within the AL [4, 113]. Therefore, glomerular groups in locusts cannot be functionally analogous to single glomeruli in other insects.

Finally, the role of LNs in the antennal lobe is to inhibit PNs according to the odor information they received from ORNs. This lateral inhibition seems to perform a gain control mechanism, which regulates the neural activity of PNs and make it invariant to changes in the concentration of odorants [168]. LNs are also divided into two classes: the first type innervates most if not all glomeruli uniformly (homoLNs) while the other innervates only a few of them (heteroLN) [3, 38, 46, 133, 212].

2.2.3 Mushroom body

The mushroom bodies (corpora pedunculata) of the insect brain are a pair of easily discernible neuropils separated from the rest of the brain by a thin sheath of glia lamellae. Mushroom bodies (MBs) are involved in learning [30, 76, 210], and receive both olfactory and visual input in most insect species [42, 211]. The basic elements of the MBs are the Kenyon cells, calyces, and lobes (see Fig.2.6). Kenyon cells (named after F. C. Kenyon, who first described them in 1896 [105]) are densely packed in a region that outlines cup-shaped neuropils called calyces [141], which receive the dendritic tree from the Kenyon cells (KCs). The axons of the KCs group together at the center of the calyx to form a vertical column of parallel fibers. This vertical column is called the stalk of the mushroom body. In this area, the axons lack side branches and spines, but at the lower end of the stalk, the axons

branch once, giving rise to two lobes pointing in roughly orthogonal directions: vertical (α-lobe) and medial (β-lobe) [171].

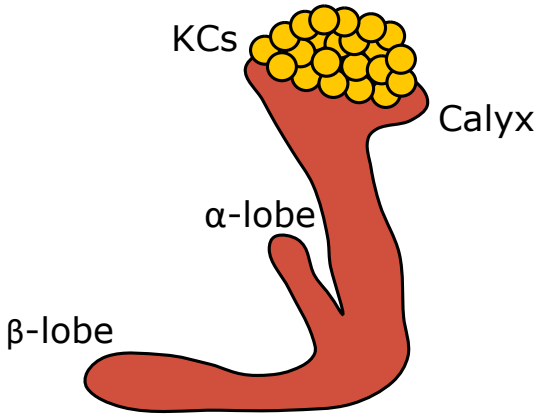


Figure 2.6: Structure of the mushroom body. The basic elements of the MBs are the Kenyon cells, calyces, and lobes. Kenyon cells are densely packed in a region that outlines cup-shaped neuropils called calyces [141], which receive the dendritic tree from the Kenyon cells (KCs). The axons of the KCs group together at the center of the calyx to form a vertical column of parallel fibers that branch into two lobes pointing in roughly orthogonal directions: α-lobe and β-lobe [171].

KCs are intrinsic neurons that present different types of neural heterogeneity in their populations [41]. Subdivisions of the Kenyon cell population have been described on the basis of the morphology of arborizations in the calyces and projections to the lobes. Other subdivisions are based on patterns of immunoreactivity, patterns of gene expression, or relative diameter of the somata. The number of Kenyon cells changes between insects [42], ranging from $2,500$ (fruit fly) to $230,000$ (cockroach).

KCs synapse with PNs in the antennal lobe through the calyx glomeruli. The connectivity pattern of KCs dendrites in calyx glomeruli is random [21, 132, 237], in contrast to the highly stereotypical pattern of connections observed for PNs in the antennal lobe glomeruli. This random pattern of connectivity for KCs suggests that one role of calyx glomeruli is to allow the conversion of the olfactory information from a limited number of channels transmitted by PNs, to a high-dimensional representation in the MB that potentially allows discrimination of a much larger palette of odor combinations [129, 132]. This is facilitated by the fan-out phase

between PNs and KCs, since the number of PNs is much lower than the number of KCs (for example, in locust, there are 830 PNs that send information to approximately 50,000 KCs), which increases the separability of odor patterns [62]. Furthermore, KCs also show a different behavior than PNs, since they remain silent most of the time, with a mean firing frequency lower than $1Hz$ [175]. Thus, we can assume that they act as coincidence detectors [6] that only produce a spike response when they receive a specific number of inputs from a subset of PNs within some time frame ($\pm 10ms$). This time frame is set by the fact that KCs activity is reset every 50 milliseconds, erasing the state they were previously in, as we will discuss in section 2.2.4.

KCs send their information to the MBs output neurons (MBONs), which are responsible for the final identification of the odorants. The association between the stimuli representation in KCs and the identity of the odorants in the MBONs is achieved through synaptic plasticity in the connections between the two types of neurons. This, in combination with the mutual inhibition observed between the MBONs, is a common mechanism for learning in neural networks [161, 169].

2.2.4 Lateral Horn

The lateral horn (LH) is the area defined by the terminal arborizations of PNs in the lateral protocerebrum [39, 82], which remains poorly understood.

In [175], the authors focused on a population of interneurons in the LH (LHIs), a subtype of LH neurons (LHNs), that have synapses with the MB and appear to be GABAergic. Since Kenyon cells (KCs) in the MB receive oscillatory waves of excitatory input directly from PNs, [175] proposed that the LHIs, which also receive this excitatory input from PNs, feed it forward, after a brief delay, as inhibition to KCs. KCs would thus receive alternating waves of excitation from PNs and inhibition from LHIs, which together define temporal integration windows that help maintain the minimal and specific odor-elicited spiking observed in KCs. The duration of the time windows between successive inhibition waves from LNs would be around 50 milliseconds.

The LH has also been proposed to encode innate olfactory preferences. Abolishing the MB in *Drosophila* was linked to deficits in olfactory learning but had

little effect on innate olfactory behaviors [32, 106]. Further, reducing olfactory input to both the MB and the LH caused deficits in innate behaviors, leading [75] to argue, by exclusion, that the LH could mediate innate, unlearned responses. Consistent with this proposal, anatomical studies in *Drosophila* showed stereotypical clustering of PN arborizations in the LH [129, 217, 237]. Further, the LH appears to provide labeled lines for processing pheromones and mediating sexual behaviors [98, 188]. However, whether the LH provides specialized pathways for specific odors processed by the general (nonpheromonal) olfactory system remains unclear because LHNs could integrate input from diverse PNs [98] and blocking synaptic transmission in the MB can cause deficits in innate behaviors [230, 241].

However, in [72], the hypothesis that LH can control the sparseness condition of KCs and encode innate olfactory preferences is rejected. Firstly, it is argued that LHIs are not GABAergic. Moreover, nearly all the GABAergic fibers observed in the MB calyx overlapped precisely with arbors of the giant GABAergic neuron (GGN) of the locust. This indicates that GGN provides by far the greatest source of inhibitory input to the KCs. Thus, of the two types of neurons (GGN and LHIs) proposed to inhibit KCs, only GGN appears to perform this role. This could also be the case for the anterior paired lateral neuron (APL) of the *Drosophila* [120, 176, 239] and the large GABAergic neurons of cockroaches [242] and honeybees [12] because of their similarity with the locust GGN [83, 172]. Finally, all classes of LHNs analyzed in [72] responded broadly to the panel of odorants. Therefore, it was inferred that these LHNs are not specialized to respond exclusively to any potentially relevant odors, discarding them as detectors of innate olfactory responses. Instead, they suggest that the LH may perform other sensory functions: representing general odor properties such as intensity, starting bilateral olfactory integration and participating in multimodal integration. These possible functions could all contribute to not innate olfactory tasks such as odor tracking.

2.3 Model of insect olfactory system

There are several models of olfactory systems [9, 11, 29, 31, 37, 89, 90, 108, 128, 139, 160, 161, 162, 247] constructed from different neuronal models: Hodgkin-Huxley [80], McCulloch-Pitts [134], Davison [31], leaky integrate-and-fire [204],

etc. These models tend to focus on the interaction of a small number of neurons due to the computational cost of simulating large neuronal models. For example, the models focused on the gain control mechanism observed in the AL [11, 24, 119, 163, 197, 247]. One of the largest neural networks that represents the main layers of the olfactory system for an expensive neuronal model (Hodgkin-Huxley neurons) is presented in [160, 161] for *Drosophila*. This model is a single-hidden-layer neural network in which the input layer corresponds to PNs in AL, the hidden layer to the KCs and the output layer to the MB output neurons (MBONs). In this model, the learning observed between KCs and MBONs is achieved by spike timing dependent plasticity rules. Furthermore, there is a fourth group of neurons which represents LHNs and performs the gain control mechanism between PNs and KCs, so that the insect olfactory system remains invariant to the concentration changes. However, although [160, 161] used a smaller system that roughly follows the statistics of *Drosophila* (100 PNs, 20 LHNs, 2500 KCs and 100 MBONs) to limit the size of the neural network, the model still has the disadvantage of a high computational cost.

These limitations are not present in the model detailed in [89, 90] due to its simplicity. Since KCs are mostly silent and act as coincidence detectors, performing a single spike followed by a reset, the McCulloch-Pitts model can be used to simulate their behavior. This type of neurons eliminates the temporal dimension in the model. Therefore, the gain control, by which the activity of the PNs remains almost constant, can be introduced directly into the patterns as a normalization of the stimuli. Finally, the learning observed in the MB is applied in this model using Hebbian rules. The combination of Hebbian learning and the winner-take-all approach in the output layer, supported by the lateral inhibition observed in MBONs, represents a biologically feasible mechanism to account for learning in neural networks [161, 169].

Because of the simplicity of this model, its reduced computational cost and the fact that it retains the most important characteristics of the insect olfactory system, we consider it an appropriate starting point to accomplish the aims of this thesis by introducing the necessary modifications and expansions on it.

2.4 Neural heterogeneity

Mathematical models of neurons and neural circuits are useful tools to analyze and understand real neural systems. However, although single-neuron models can have a high level of biophysical detail, most of these details are usually removed when modeling larger systems for the sake of simplicity. This is the case of the natural variability found in the biophysical properties of neurons [8, 73, 99, 180], even among same-class neurons, which is averaged out in most theoretical and computational modeling studies. But experimental studies have shown that neural heterogeneity does have non-trivial effects on information processing mechanisms. For example, in the electroreceptors of the South American weakly electric fish *Apteronotus leptorhync*, it can be observed that neural heterogeneity affects their bursting coding in vivo [7] and the envelope and temporal coding in peripheral neuronal populations [195]. Because of this, new models that take into account the neural heterogeneity of biological systems are necessary to understand neural coding.

This has led to to several studies that analyze the impact of neural heterogeneity in neuronal correlations [26, 245] and synchronization [16, 34, 66, 67, 125, 136, 137, 156, 166, 215, 216, 233], detection of weak signals [173, 219] and the different types of neural coding [26, 91, 136, 195]. Other components of the biologycal systems are also prone to show heterogeneity that have been traditionally ignored in models are neuronal threshold [48, 77, 85, 86, 96, 174, 234, 244] and responses to stimuli [28, 51, 52, 182, 183, 224, 247]. This is the reason why, in this thesis, we are going to analyze the role of heterogeneity in both thresholds and neuronal responses and how it can influence the processing of information.

Neural threshold is the potential that must be reached in the cellular membrane of the neuron in order to produce a neural spike. Although it was initially thought that neural threshold had a fixed value, many studies have shown that neural thresholds are dynamic and adapt to the input signal that arrives at the neuron. Hence, the existence of neural thresholds dynamics linked to each neuron in a biological system could constitute a way of neural heterogeneity and could have important repercussions in the way information is processed. Although the neural threshold has not been vastly studied yet, several studies have associated

its dynamics with different computational capacities found in very different biological systems. Between these capacities are the selectivity of neurons to certain attributes of the stimuli they receive [40, 234], as a mechanism to achieve short term memory [77, 96], filtering the input signal [48, 85], synchronization in neural populations and ensuring a robust [49, 86] and energetic efficient codification of the information [100, 244]. In the case of locust, in section 2.2.3 it was said that KCs behave as coincidence detectors that only produce a response when the input they receive is strong enough, which means when the input reaches a certain neural threshold. Hence, the neural threshold of KCs seems to be a powerful component of the system, capable of regulating the intensity of the response of each neuron and the overall activity in the population. The existence of KCs with heterogeneous threshold values in this system could introduce and regulate interesting computational properties which we will explore through a computational model in the following chapters.

In the case of specialist and generalist neurons (observed in the olfactory system of insects [28, 175, 182, 183, 247], the gustatory neurons from rat geniculate ganglion [51, 52, 126] and deterrent chemoreceptors from *Pieris* caterpillars [224]), their populations show heterogeneity in their responses to stimuli. Specialist neurons are selective responding to stimuli, while generalists code for multiple stimuli [28]. The role of both classes of neurons in the olfactory system is still unclear [28, 102]. The common hypothesis is that specialist neurons are crucial for discrimination, while generalist neurons play a key role in extracting and discovering common features [236]. A hypothesis that we have analyzed by regulating the ratio of specialist and generalist neurons in the KCs and estimating its impact on the odor discrimination capability of the system.

Finally, the gain control mechanism observed in AL is performed by the heterogeneous inhibition of LNs. LNs are divided into two types of neural populations: homoLNs, which innervate most if not all glomeruli uniformly, and heteroLNs, which innervate only a few of them [65]. The aim of this study is to simulate this gain control mechanism and analyze its ability to modulate the stimulus signal compared to normalization. The fact that the inhibition of LNs is heterogeneous may be the real reason why this gain control mechanism is present in AL, since inhibitory heterogeneity and gain control seem to be related [137].

CHAPTER

3

Materials and methods

This section presents all the materials and methods used to understand how odor information is processed in the insect olfactory system and, especially, the role of neuronal diversity and gain control mechanism on it and their effects on the odor pattern recognition.

In section 1, we present the computational model used to simulate the insect olfactory system. This model is based on the one described in [89], although with variations depending on the neuronal diversity that we want to study at each moment. We will detail the network and neural model used, as well as the mechanisms for the odor classification in the insect olfactory system: the learning algorithms used to simulate the synaptic plasticity of KCs-MBONs connections [22, 23] and the lateral inhibition observed in MBONs [199].

Section 2 describes the methods implemented to analyze the role of heterogeneity in neural thresholds, stimuli response observed in KCs and inhibition responsible for gain control mechanism in AL.

Finally, in section 3, we will introduce the input patterns, which represent the neural activity in AL. First, we will describe the patterns developed in this thesis: primary and Gaussian patterns. Primary patterns were used only in the first stages of the thesis, since their purpose was to allow a first contact with the subject under study. On the other hand, Gaussian patterns emerged as a way to simulate the

activity observed in the glomeruli for different odors [53, 208]. These odor patterns also allow us to analyze how the overlap degree between them (which determines the complexity level of the pattern recognition problem) affects the subsequent classification capacity of the simulated olfactory system. To generalize the results obtained by the previous patterns, we have used two external databases: MNIST digits [115] and chemical data obtained by electronic noses [184, 225]. The first ones are data widely used in pattern recognition problems, while the second is closer to the chemical nature of the odorants that we want to understand how they are processed. These three last types of patterns are the most widely used in this thesis, generally together in order to compare the different results obtained.

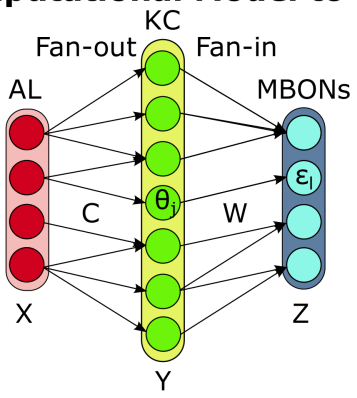
3.1 Computational model of insect olfactory system

To represent the olfactory system of insects we used a computational model that takes the most essential parts of this system, AL and MB, which are involved in the function of encoding and classifying the odor stimuli received by ORNs. The classification process is divided in two phases: one based on increasing the separability between the patterns, performed by KCs, and the classification in itself by the MBONs. We used a single-hidden-layer neural network (SLN) to model these parts of the olfactory system where the input layer represents the AL, the hidden layer the KCs and the output layer the MBONs.

However, this model (Fig.3.1), based on the one described in [89, 90], has undergone variations throughout this thesis depending on the type of neuronal diversity that we want to study at each moment.

For the study on neural thresholds, we started with a simple version of this model (see panel (a) in Fig. 3.1). The gain control mechanism in the input layer was performed by the normalization of patterns, the selection of homogeneous and heterogeneous thresholds was performed by exhaustive search and the learning of W matrix (which represents the synaptic plasticity between KCs and MBONs) was performed by Hebbian learning. The results obtained through this computational model can be seen in section 4.1.1, as well as in the following publications [144, 146]. In order to improve the selection of homogeneous and heterogeneous thresholds, we developed a gradient descent algorithm for the neural thresholds and

a Computational Model to study Neural Threshold:



Gain control in AL:
Normalization

Neural model:

$y_j = \varphi(\sum_{i=1}^{N_{AL}} c_{ji}x_i - \theta_j)$

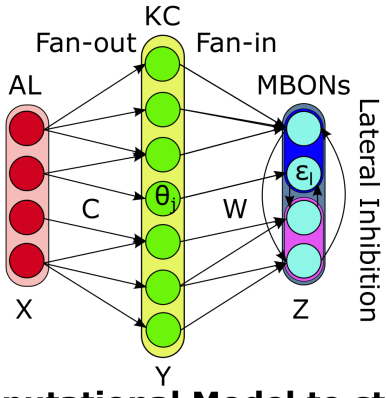
$z_l = \varphi\left(\sum_{j=1}^{N_{KC}} w_{lj}y_j - \varepsilon_l\right)$

Learning algorithm in W:

· Hebbian
(φ step function)

· Gradient Descent + ELM
(φ sigmoid function)

b Computational Model to study Specialists/Generalists:



Gain control in AL:
Normalization

Neural model:

$y_j = \varphi(\sum_{i=1}^{N_{AL}} c_{ji}x_i - \theta_j)$

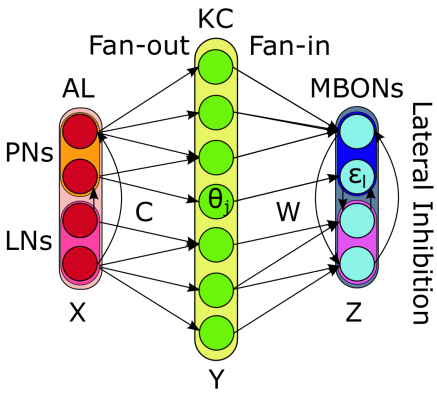
$z_l = \varphi\left(\sum_{j=1}^{N_{KC}} w_{lj}y_j - Inhib - \varepsilon_l\right)$

$Inhib = \frac{1}{N_{class}} \sum_{k=1}^{N_{class}} \sum_{j=1}^{N_{KC}} w_{kj}y_j$

Learning algorithm in W:

· Hebbian

c Computational Model to study Gain Control Mechanism:



Gain control in AL:

· Homogeneous

$PN_j = PN_j - PN_j\beta\left(\sum_{i=1}^{N_{LN}} LN_i - \delta\right)$

· Heterogeneous:

$PN_p^{GL_i} = PN_p^{GL_i} - \alpha Inhib_{homoLN} - \beta Inhib_{heteroLN}$

$Inhib_{homoLN} = \left(\sum_{j\neq i}^{N_{GL}} \sum_{k=1}^{N_{homoLN^{GL_j}}} homoLN_k^{GL_j}\right)$

$Inhib_{heteroLN} = \sum_{l=1}^{N_{heteroLN^{GL_i}}} heteroLN_l^{GL_i}$

Neural model:

$y_j = \varphi(\sum_{i=1}^{N_{AL}} c_{ji}x_i - \theta_j)$

$z_l = \varphi\left(\sum_{j=1}^{N_{KC}} w_{lj}y_j - Inhib - \varepsilon_l\right)$

$Inhib = \frac{1}{N_{class}} \sum_{k=1}^{N_{class}} \sum_{j=1}^{N_{KC}} w_{kj}y_j$

Learning algorithm in W:

· Hebbian

Figure 3.1: Models used in different studies on neuronal diversity. All are SLNs with AL as input (X), KCs as hidden layer (Y) and MBONs as output (Z). The connectivity matrices are C and W and θ_j and ε_l the thresholds of Y and Z.

the weight of W matrix. Despite the good results obtained (see section 4.1.2), the computational cost of this learning algorithm was too high due to its complexity.

In the case of the study of specialist and generalist neurons in KCs, differentiated in the previous study due to their neural threshold variability, we decided to introduce some changes in the model in order to make it faster (see panel (b) in Fig. 3.1). Instead of selecting the neural thresholds for improving the odor classification, we selected threshold values for generating neurons with different degrees of response to odorants (different types of generalists and specialists). On the other hand, we changed our unsupervised Hebbian learning by a supervised one and implemented the lateral inhibition of MBONS to facilitate the clustering of the output. The results obtained through this computational model can be seen in sections 4.2-4.4 and in [145, 147, 148, 149].

Finally, using the previous computational model as a basis, we introduced in the input layer the gain control mechanisms described in section 3.2.3 (see panel (c) in Fig. 3.1), so that we could analyze the effects of these mechanisms on the pattern recognition. Their results are shown in section 4.5 and in [150].

3.1.1 Neural network model

The neural network of the computational model used is an SLN where AL represents the input (X), KCs the hidden layer (Y) and MBONs the output (Z).

The dimension of the input layer depends on the number of attributes of the patterns (see section 3.3), while the number of output neurons is determined by the number of classes. In case the output layer has lateral inhibition, we have 10 neurons for each pattern class, otherwise, we have a single neuron per class. Finally, in the case of the hidden layer, the number of neurons depends on the insect whose system we are trying to simulate. During this thesis, we used the dimensions of two types of insects: locust and *Drosophila*. In the case of the locust, the insect that we generally had simulated, the PNs-KCs ratio is $1:50$, always taking the total number of KCs in the locust $(50,000)$ [175, 243] when we used with MNIST digits (784 input neurons, as mentioned section 3.3.3). On the other hand, the PNs-KCs ratio of the *Drosophila* is $1:10$ and the KCs number is $2,500$ [153, 202].

The connectivity matrices that link the layers of the neural network are C and W. The C matrix (PNs-KCs connections) is determined randomly by independent Bernoulli processes with probability p_c for each existing connection and $1 - p_c$ for each lack of this [62, 89, 90]. The reason for this non-specific connectivity matrix is because there is no reproducibility in the neural connections between AL and MB across individuals of the same species [21, 132, 237]. In the case of W matrix (KCs-MBONs connections), is initialized randomly but subsequently updated using Hebbian learning [36, 89, 90] or our gradient descent algorithm (see section B.2), because MB is involved in memory formation and storage [36, 135, 246]. In the first studies, we initialized the matrix W by a Bernoulli process like the one used for matrix C, using a probability p_w [144, 146]. However, we subsequently varied this type of initialization to a random matrix of natural numbers for improving the Hebbian learning used [145, 147, 148, 149, 150].

Finally, for thresholds values of hidden and output layers (θ_j and ε_l respectively), we opted for using different thresholds for KCs (heterogeneous thresholds) and the same threshold for all MBONs(homogeneous threshold) for most of the thesis, after comparing between these two types of thresholds. This comparison showed (see section 4.1) that using heterogeneous thresholds in the hidden and output layers improved the odor classification. However, once lateral inhibition was implemented in the output layer (see section 3.1.3), no additional mechanism was necessary to help to odor discrimination.

3.1.2 Neural model

The KC neurons of the MB display very low activity [175]. They are inactive most of the time, with a mean firing frequency lower than 1 Hz. But when they are activated, their neuronal response is produced by the coincidence of concurrent spikes followed by a reset. Bearing in mind this behavior, we chose the McCulloch-Pitts model for simulating all neurons of the hidden and output layers. In the case of KCs, due to the reasons given, in the case of MBONs, because we only need the spatial information of output patterns to compare them with the target ones. This neuron model uses the threshold step function as the activation function. Therefore,

the activation of KCs in the hidden layer and MBONs in the output is given by [144, 145, 146]:

$$y_j = \varphi(\sum_{i=1}^{N_{AL}} c_{ji}x_i - \theta_j), \;\; j = 1, \ldots, N_{KC},$$
$$z_l = \varphi(\sum_{j=1}^{N_{KC}} w_{lj}y_j - \varepsilon_l), \;\; l = 1, \ldots, N_{MBON}, \tag{3.1}$$

where x_i, y_j and z_l are activation states for a input, hidden and output neuron respectively, c_{ji} and w_{lj} are weights linking two neurons, θ_j and ε_l are thresholds for the hidden and output neuron, respectively, and φ is the Heaviside activation function (see network model in Fig.3.1). The Heaviside activation function φ is 0 when its argument is equal or inferior to 0 and 1 otherwise.

3.1.3 Lateral inhibition

The equation for the output neurons changes when we introduce lateral inhibition [199] to MBONs [147]. In this case, we have a population of neurons in the output layer for each pattern class. Each of these groups of neurons receives the inhibition of the other ones [10]. Therefore, we use the winner-take-all concept [178] as activation function for the groups of MBONs as follows:

$$z_l = \varphi\left(\sum_{j=1}^{N_{KC}} w_{lj}y_j - \frac{1}{N_{class}}\sum_{k=1}^{N_{class}}\sum_{j=1}^{N_{KC}} w_{kj}y_j - \varepsilon_l\right), \; l = 1, \ldots, N_{class}, \tag{3.2}$$

where z_l represents the activation state of a group of MBONs, which are specialized in a certain pattern. We have N_{class} patterns and, therefore, the same number of groups of MBONs. The weight that links two neurons is w_{lj} and the neural threshold for the output layer is ε_l. The Heaviside activation function φ is only 1 for the winner MBONs group.

3.1.4 Learning algorithms

As mentioned above, the connectivity matrix W linking the KC and MBONs undergoes associative learning. This learning can be simulated using Hebbian

learning [10, 36, 143]. However, we can also use gradient descent or the ELM algorithm [87, 185] to perform this learning. Although these two last methods are not bio-inspired, they are going to be introduced in section 3.2.1 as an option to calculate the neural thresholds.

3.1.4.1 **Hebbian learning**

Hebbian learning is the classical choice for local synaptic changes [74]. It allows strengthening or weakening the connections of a connectivity matrix. At the beginning of this study [144, 146], the strengthening and weakening of the weights were absolute, the connection was created or destroyed completely, and the learning was unsupervised, as shown below [90]:

$$w_{lj}(n+1) = H(y_j, z_l, w_{lj}(n)),$$

$$H(1,1,w_{lj}(n)) = \begin{cases} P(H(1,1,w_{lj}(n)) = 1) = p_+ \\ P(H(1,1,w_{lj}(n)) = w_{lj}(n)) = 1 - p_+ \end{cases}$$

$$H(0,1,w_{lj}(n)) = \begin{cases} P(H(0,1,w_{lj}(n)) = 0) = p_- \\ P(H(0,1,w_{lj}(n)) = w_{lj}(n)) = 1 - p_- \end{cases} \tag{3.3}$$

$$H(1,0,w_{lj}(n)) = w_{lj}(n),$$

$$H(0,0,w_{lj}(n)) = w_{lj}(n),$$

where the future connection state $w_{lj}(n+1)$ is determined by a function H. This function $H(y_j, z_l, w_{lj}(n))$ depends on the hidden layer neuron y_j, the output neuron z_l and the current connection state $w_{lj}(n)$.

If the output neuron z_l fires, the connection state depends on the hidden layer in the following ways:

- If the hidden layer neuron y_j has fired, then the connection between these neurons is created with a probability p_+.

- If the hidden layer neuron y_j has not fired, then the connection between these neurons is destroyed with a probability p_-.

This learning was replaced later by a supervised one, which changed the output neuron z_l by the target set for this output neuron, t_l [145, 147, 148, 149, 150]. Finally, we introduced a progressive strengthening and weakening of the weights and a negative reinforcement of them [89]. Therefore, the final Hebbian learning and negative reinforcement rules are the following:

$$w_{lj}(n+1) = H(y_j, z_l, t_l, w_{lj}(n)),$$

$$H(1, z_l, 1, w_{lj}(n)) = \begin{cases} P(H(1, z_l, 1, w_{lj}(n)) = w_{lj}(n) + 1) = p_+ \\ P(H(1, z_l, 1, w_{lj}(n)) = w_{lj}(n)) = 1 - p_+ \end{cases}$$

$$H(0, z_l, 1, w_{lj}(n)) = \begin{cases} P(H(0, z_l, 1, w_{lj}(n)) = [w_{lj}(n) - 1]_+) = p_- \\ P(H(0, z_l, 1, w_{lj}(n)) = w_{lj}(n)) = 1 - p_- \end{cases} \quad (3.4)$$

$$H(1, 1, 0, w_{lj}(n)) = \begin{cases} P(H(1, 1, 0, w_{lj}(n)) = [w_{lj}(n) - 1]_+) = p_+ \\ P(H(1, 1, 0, w_{lj}(n)) = w_{lj}(n)) = 1 - p_+ \end{cases}$$

$$H(1, 0, 0, w_{lj}(n)) = w_{lj}(n),$$

$$H(0, z_l, 0, w_{lj}(n)) = w_{lj}(n),$$

where the future connection state $w_{lj}(n+1)$ is determined by a function H. This function $H(y_j, z_l, t_l, w_{lj}(n))$ depends on the hidden layer neuron y_j, the current value for the output neuron z_l, the target for this output layer neuron t_l and the current connection state $w_{lj}(n)$. On the other hand, $[x]_+$ denotes that x cannot be negative, since the connections that start from KCs are only excitatory [175]. In case of x takes a negative value, its value becomes 0.

As a summary, if the target neuron t_l fires, the connection state depends on the hidden layer in the following ways:

- If the hidden layer neuron y_j has fired, then the connection between these neurons is reinforced with a probability p_+.
- If the hidden layer neuron y_j has not fired, then the connection between these neurons is weakened with a probability p_-.

Finally, if the target for the output neuron t_l does not fire, the connection state depends on the current value of output neuron z_l and the hidden layer neuron y_j. If the hidden and output layer neurons have fired but the target of output neuron did not fire, then the connection between these neurons is weakened with a probability p_+ (negative reinforcement). Otherwise, the connection state is not changed.

3.1.4.2 **Gradient descent**

Another way to learn the weights of the matrix W is by gradient descent. This method modifies the value of the weights in order to reduce the error between the target, T, and the output of the neural network, Z, as follows:

$$E = |T - Z|^2 \tag{3.5}$$

$$w_{nl} = w_{nl} - \eta_w \frac{\partial E}{\partial w_{nl}} \tag{3.6}$$

where w_{nl} is the weight between the hidden layer neuron n and the output layer neuron l and η_w is the learning rate for the weights of matrix W. Since the Heaviside step function is not derivable, we use a sigmoid function in the output neurons to perform the gradient descent:

$$z_{kp} = \frac{1}{\left(1 + \exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right))\right)} \tag{3.7}$$

where y_{jp} and z_{kp} are the neural activities for the neurons of the hidden and output layers (j and k respectively) for a pattern p. The neurons y_j and z_k are linked by the weight w_{jk} and ε_k is the threshold of z_k. Finally, γ is the coefficient that defines the sigmoid slope. Taking this into account, we calculate $\frac{\partial E}{\partial w_{nl}}$ as follows:

$$\frac{\partial E}{\partial w_{nl}} = 2\alpha \sum_{p=1}^{N_P} (z_{lp} - t_{lp}) \frac{\exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l\right))}{\left(1 + \exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l\right))\right)^2} \gamma y_{np} \tag{3.8}$$

This gradient descent algorithm for calculating weights is the classic method. In this thesis we have proposed a more complex one (see equation 3.12) that we can use, for example, to control the activity of the hidden layer (KCs) and to achieve the sparse condition of these neurons [175].

3.1.4.3 ELM Algorithm

An alternative to the gradient descent is the ELM algorithm [87, 185]. This method optimize the calculation of the weights of matrix W using the Moore-Penrose generalized inverse (†) as follows:

$$W = Y^{\dagger}T \tag{3.9}$$

where T is the target output matrix, Y is the hidden layer matrix and W are the optimal weights that allow us to reach these targets. Therefore, this method implies an immediate adjustment of the weights, instead of the progressively one used by Hebbian learning and gradient descent.

3.2 Neural heterogeneity

Biological experiments have shown the relevance of neural heterogeneity in natural systems [7, 81, 130, 170, 195]. Hence, we want to study its role in the insect olfactory system based on its effects on odor classification. In particular, we focus on analyzing the three kinds of neuronal diversities observed in neural thresholds, the neural populations of specialists and generalists in KC and the gain control mechanism of AL. In this section, we define the methods used to analyze the implications in pattern recognition of these types of heterogeneity.

3.2.1 Neural thresholds

Typical models of the olfactory system show very little variability in the excitability of the neurons, implemented by fixed neural thresholds. However, neural thresholds are not fixed in biological systems, as it can be observed in ORNs [2] and KCs [175] of the insect olfactory system. There are different hypothesis about the role of this threshold variability [123]: differentiation of neural populations

[177, 218], feature selectivity [40, 234], energy efficiency [100, 244], short-term memory [77, 96], robust coding [49, 86], filtering and synchronization [48, 85]. Additionally, applied research on artificial noses determined that using heterogeneous detection thresholds for different odorants can improve gas discrimination [35, 47]. For these reasons, we considered that it could be interesting to study the application of heterogeneity in neural thresholds for solving pattern recognition problems.

To analyze the advantages of using different thresholds for each neuron of the KC layer (heterogeneous thresholds) instead of using a unique threshold for all of them (homogeneous ones) for pattern recognition problems, we used two methods for threshold selection: exhaustive search and gradient descent algorithm.

3.2.1.1 Exhaustive search

The threshold selection that we performed by exhaustive search is based on the concept of limit threshold. A limit threshold is the total stimulation received by a neuron for a given odorant and, therefore, the minimum threshold value which prevents that the neuron fires for that odorant (see Fig. 3.2). This limit threshold was calculated for each neuron and each odorant as follows:

$$\theta_j^o = \sum_{i=1}^{N_{AL}} c_{ji} x_i^o, \qquad \varepsilon_l^o = \sum_{j=1}^{N_{KC}} c_{lj} y_j^o \tag{3.10}$$

where neuron j spikes $\forall \theta_j, 0 \leq \theta_j < \theta_j^o$, and neuron l spikes $\forall \varepsilon_l, 0 \leq \varepsilon_l < \varepsilon_l^o$. Being θ_j^o the limit threshold for a KC (j) and an odorant (o) and ε_l^o the limit threshold for a MBON (l) and an odorant (o). These thresholds were calculated only one time in the classification process before Hebbian learning is applied. The limit threshold matrices store all limit thresholds of KCs and MBONs and have dimension $N_{KC} \times N_{ODOR}$ and $N_{MBON} \times N_{ODOR}$ respectively:

$$\theta_j^o = \begin{pmatrix} \theta_1^1 & \cdots & \theta_1^{N_{ODOR}} \\ \vdots & \ddots & \vdots \\ \theta_{N_{KC}}^1 & \cdots & \theta_{N_{KC}}^{N_{ODOR}} \end{pmatrix} \qquad \varepsilon_l^o = \begin{pmatrix} \varepsilon_1^1 & \cdots & \varepsilon_1^{N_{ODOR}} \\ \vdots & \ddots & \vdots \\ \varepsilon_{N_{OutN}}^1 & \cdots & \varepsilon_{N_{OutN}}^{N_{ODOR}} \end{pmatrix} \tag{3.11}$$

The purpose of these matrices is to collect all possible thresholds for each layer (for all neurons and odorants) in order to subsequently select the homogeneous and heterogeneous thresholds used in our model.

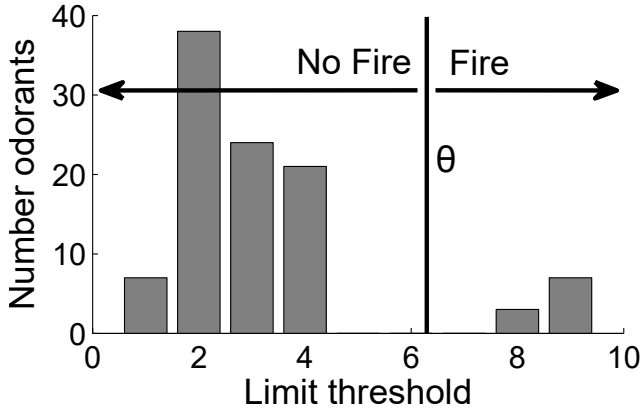


Figure 3.2: Example of limit thresholds distribution for a neuron. This figure displays how the limit thresholds of a neuron are distributed (based on the different stimulations received) for a total of 100 odorants. The minimum threshold value is 1 and the maximum value is 9. In this example, we established that the neural threshold (θ) is 6, so all the odorants that stimulate the neuron above this value will make that the neuron responds to them. We can reduce or increase the θ value depending on whether we want to increase or reduce the number of odorants to which the neuron responds.

In the case of homogeneous thresholds (see algorithm 1 in section B.1.1), the minimum and maximum values for these matrices determine the range of selected thresholds. We calculated the classification error for each possible combination of threshold values in the hidden and output layer to find the one that minimizes this error. This minimum classification error, for homogeneous thresholds, is what we used to compare the results obtained by heterogeneous thresholds [144, 146].

In the case of heterogeneous thresholds, we used the distribution of limit thresholds for each neuron for two different purposes:

- To find the threshold values, for the hidden and output layers, that improves the classification error and to compare them with the homogeneous ones [144, 146].
- To increase the neural sensitivity in KCs to analyze which ratios of specialist and generalist neurons obtain the minimum classification error [145, 147, 148, 149, 150].

For the comparison between homogeneous and heterogeneous thresholds, we selected the heterogeneous ones by two different methods:

- Firstly, we were interested in a simple method for calculating the heterogeneous thresholds without performing all possible combinations of limit thresholds ($N_{KC}^{N_{ODOR}}$ and $N_{MBON}^{N_{ODOR}}$). Because neural thresholds prevent neurons from firing if they do not receive enough input intensity, we implemented a method that prevents neurons from firing for a certain percentage of odorants (see algorithm 2 in section B.1.2.1) using the input intensity that they generate in each neuron (limit threshold). Thus, using percentages from 0% to 100% for the neurons of hidden and output layers, we start from a configuration in which heterogeneous thresholds allow firing all neurons for all odorants (0%) until the final configuration, where no neuron can fire for any odorant (100%). The total number of combinations for this method is 101^2 (see lines 5 and 9 of algorithm 2).

- The previous method for obtaining heterogeneous thresholds showed good results for the primary patterns [144], but not for other pattern sets such as electronic noses data. To solve this problem, we decided to increase the degree of heterogeneity of thresholds, because even if all neurons have a different threshold value, the percentage of odorants for which they fire is homogeneous. Then, we analyzed the different limit threshold distributions for each neuron and we observed that we could differentiate two types of neurons: some that responded to all odorants with the same or similar intensity and others with a greater range of response intensities. This behavior is consistent with the populations of generalist and specialists neurons in KCs [175]. Hence, taking this into account, we decided to classify the neurons in these two types using the variance of their limit thresholds distribution. Because it was suggested that only specialist neurons are necessary for odor discrimination, we assigned the maximum limit threshold to generalist neurons with the objective of "pruning" them. On the other hand, in the case of specialist neurons, we assumed that, due to the great variability of these neurons, their distribution could be bimodal. This entails that there is a minimum value between the two different modes. Hence, we developed a method that tries to find this minimum value, making the neuron responds only to a few odorants (specialist) since the second mode is usually inferior to the first

one. The percentage of odorants to which a neuron stops responding to by this method is called the specialization coefficient α. However, given the impossibility of determining which variance of thresholds divides specialist and generalist neurons, our algorithm initially considered all neurons as specialists and progressively considered them all as generalists (covering all the integer values from 0% to 100%, see Fig. 4.4). This method (see algorithm 3 in section B.1.2.2) had some limitations, but in many cases it allowed us to achieve the proper thresholds [145, 146] (Fig. 3.3).

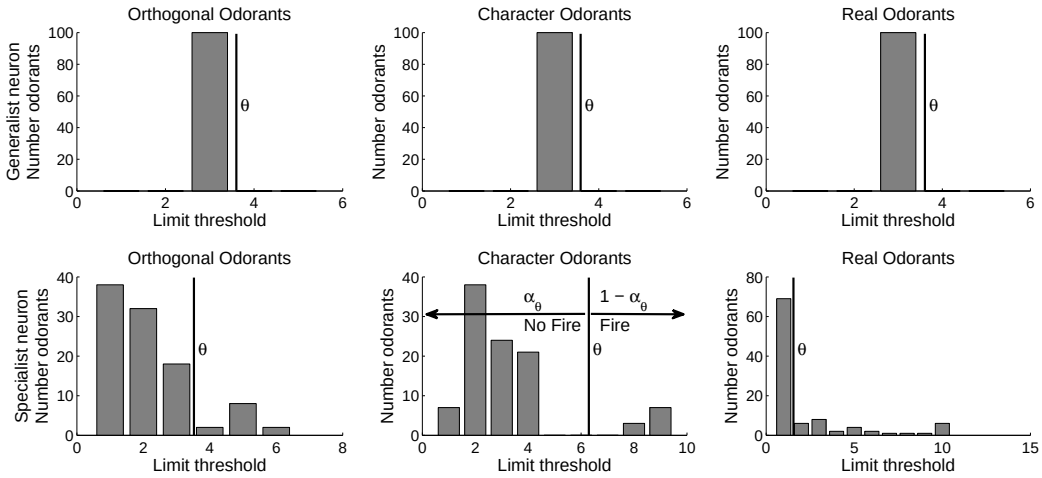


Figure 3.3: Examples of limit thresholds distributions for generalist and specialist neurons. For a generalist neuron, the neural threshold (θ) is the maximum limit threshold and, thus, this neuron cannot respond to any of the training patterns. For a specialist neuron, we search the minimum value between the two modes of the bimodal distribution. This minimum value establishes a coefficient of specialization of the neuron (α_θ).

With the aim of increasing the neural sensitivity in KCs in order to study the population of specialist and generalist neurons, we decided to group the KCs randomly in 100 sets. Then, using the first method for selecting heterogeneous thresholds, we made that the neurons of each set could only fire (be sensitive) for 1%, 2%, ... or 100% of patterns. This final method (see algorithm 4 in section B.1.2.3) for threshold selection was used for the rest of the thesis [147, 148, 149, 150].

3.2.1.2 Gradient descent

Another method of selecting optimal thresholds for the neural network is by gradient descent. Initially, we proposed an equation with four terms: the first one to calculate the optimal neural thresholds of KCs based on the ELM algorithm [87], the second one to reduce the classification error, the third one to reduce the neural activity of KCs (given the sparseness condition of these in nature [175]) and the fourth one to regulate the weights of the matrix W:

$$E = \alpha \left|Y^* - Y\right|^2 + \beta \left|T - Z\right|^2 + \lambda \left|Y^*\right| + \mu \frac{|W|^2}{2} \tag{3.12}$$

where Y^* is the target hidden layer matrix, Y is the hidden layer matrix, T is the target output matrix, Z is the output matrix and W is the connectivity matrix between Y and Z. Every term is weighted according to their coefficients, α for threshold selection, β for classification error, λ for the sparse neural activity and μ for weight regulation. Because the Heaviside function used previously in neurons of the hidden and output layer is not derivable, we used the sigmoid function as activation function instead (matrices Y and Z). On the other hand, matrix Y^* is obtained by using the ELM algorithm (see equation 3.9). Because of this, the equations for the activations of these neurons are:

$$y_{jp} = \frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)}\right)} \tag{3.13}$$

$$z_{kp} = \frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)} \tag{3.14}$$

$$y^*_{jp} = \sum_{k=1}^{N_{OUT}} w^{\dagger}_{kj} \left(\frac{-log\left(\frac{1}{t_{kp}} - 1\right)}{\gamma} + \varepsilon_k \right) \tag{3.15}$$

where x_{ip}, y_{jp} and z_{kp} are the neural activities for the neurons of the input, hidden and output layers (i, j and k respectively) for an input pattern (or stimulus) p. The targets of a pattern p for the hidden and output layers are y^*_{jp} and t_{kp}. The weight c_{ij} links the neurons x_i and y_j and the weight w_{jk} connects the neurons y_j and z_k.

θ_j is the threshold of y_j and ε_k is the one for z_k. Finally, γ is the coefficient that defines the sigmoid slope.

Hence, the update of the threshold values for the hidden and output layer (θ and ε) and the weights of the matrix W are calculated as follows (see mathematical reasoning in section B.2.1):

$$\theta_n = \theta_n - \eta_\theta \frac{\partial E}{\partial \theta_n} \tag{3.16}$$

$$\begin{aligned}\frac{\partial E}{\partial \theta_n} = \sum_{p=1}^{N_P} \Bigg[& 2\alpha \left(y^*_{np} - y_{np}\right) \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\gamma \\ & + 2\beta \sum_{k=1}^{N_{OUT}} \left(t_{kp} - z_{kp}\right) \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk}y_{jp}-\varepsilon_k\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk}y_{jp}-\varepsilon_k\right)} w_{nk} \\ & \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\gamma^2 \Bigg]\end{aligned} \tag{3.17}$$

$$\varepsilon_l = \varepsilon_l - \eta_\varepsilon \frac{\partial E}{\partial \varepsilon_l} \tag{3.18}$$

$$\begin{aligned}\frac{\partial E}{\partial \varepsilon_l} = \sum_{p=1}^{N_P} \Bigg[& 2\alpha \sum_{j=1}^{N_{KC}} \left(y^*_{jp} - y_{jp}\right) w^\dagger_{lj} + 2\beta \left(t_{lp} - z_{lp}\right) \\ & \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\gamma \Bigg] + \lambda N_P \sum_{j=1}^{N_{KC}} w^\dagger_{lj}\end{aligned} \tag{3.19}$$

$$w_{nl} = w_{nl} - \eta_w \frac{\partial E}{\partial w_{nl}} \tag{3.20}$$

$$\frac{\partial E}{\partial w_{nl}} = \sum_{p=1}^{N_P} \left[\sum_{j=1}^{N_{KC}} \left(2\alpha \left(y_{jp} * -y_{jp}\right) + \lambda\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right) \right.$$
$$\left[\left(w^T w\right)^{-1} \left(-J\left(n,l\right) \left(w^T w\right)^{-1} w^T + I\left(n,l\right)\right) \right]_{kj} + \beta \sum_{p=1}^{N_P} 2\left(z_{lp} - t_{lp}\right)$$
$$\left. \frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l\right)} \gamma y_{np} \right] + \mu w_{nl} \tag{3.21}$$

where n and l are identifiers for neurons of the hidden and output layers and η_θ, η_ε and η_w are the learning rates for the neural thresholds and the weights of the matrix W.

However, the proposed equation 3.12 had several problems. In the first place, because of the large number of terms and the difficulty for calibrating their associated constants ($\alpha, \beta, \lambda, \mu, \gamma, \eta_\theta, \eta_\varepsilon, \eta_w$), the algorithm could not escape from local minima. In addition, the complex calculation of derivatives, especially because of the target of the hidden layer (Y^*), caused a high computational cost. Finally, the difficulty for balancing the influence of the terms in the equation entailed that certain terms (such as the regulation of weights, fourth term) prevailed over the rest. Thus, we decided to simplify the previous equation as follows:

$$E = \alpha \left|T - Z\right|^2 + \beta \left|Y\right| \tag{3.22}$$

where T is the target output matrix, Z is the output matrix and Y is the hidden layer matrix. Both goals are weighted according to their coefficients, α for the classification error and β for the sparse neural activity. We are currently refining this learning, whose preliminary results can be found in [123, 124].

3.2.2 Generalist and specialist neurons

As shown in the previous section, the differentiation between specialist and generalist neurons is useful for the threshold selection. However, although it is known that specialist neurons respond selectively to stimuli, while generalists code multiple stimuli [28], their role in the olfactory system is still unclear [28, 102].

The common hypothesis is that specialist neurons are crucial for discrimination, while generalist neurons play a key role in extracting and discovering common features [236]. To analyze for ourselves this heterogeneity in neural responses to stimuli (observed in the olfactory system of insects [28, 175, 182, 183, 247], the gustatory neurons from rat geniculate ganglion [51, 52, 126] and deterrent chemoreceptors from *Pieris* caterpillars [224]) we focused on the populations of these neurons observed in KCs [175]. The reasons for studying specialists and generalists populations on KCs and not others, apart from the biological record of the existence of these two kinds of neurons in KCs, are the following:

- KCs represent the last stage of feature extraction before the classification process. Therefore, the odor information that these neurons provide is crucial for success in pattern recognition. This means that we can better observe the effects of altering the populations of these kinds of neurons in the odor classification.

- On the other hand, the large number of KCs (50,000 in locust [22, 111, 116, 175]) guarantees a sufficient number of specialist and generalist neurons for our experiments.

3.2.2.1 **Neural sensitivity estimation**

In order to classify KCs as specialist or generalist neurons, we use the concept of neural sensitivity. Neural sensitivity is the number of different stimuli for which a certain neuron responds [182, 183]. Hence, we can define specialist neurons as those with low sensitivity and generalist neurons as those whose sensitivity is high. To calculate the neural sensitivity, we need to define the minimum percentage of response to an odorant a neuron can be considered sensitive to. To estimate the suitable response degree, we made that the neurons have to respond to $20/40/60/80\%$ of patterns for a specific odorant to be considered sensitive to it (see Fig. 3.4).

As we can see in Fig. 3.4, when the response threshold is higher, the number of neurons that respond to no or few pattern classes increases, while the number of those that respond to all or many of them decreases. Since a neuron should be considered sensitive to a pattern class when it responds to it in most cases, we consider a response threshold of 80% as a desirable response percentage. This percentage

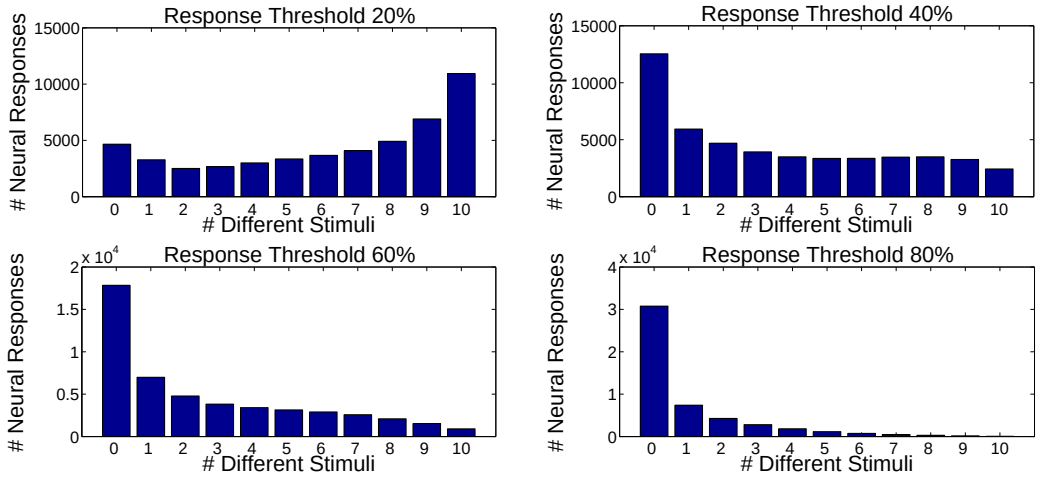


Figure 3.4: Neural sensitivity depending on the response threshold. Response threshold is the percentage of odorants from a class that a neuron needs to respond to be considered sensitive to it. When the response threshold rises, the neurons that respond to few stimuli increase in number compared to those that respond to many of them.

also entails that there are a greater number of neurons with low sensitivity, which is consistent with the neural sensitivity distribution obtained in real measurements of KCs (see Fig.3 of [182]).

3.2.2.2 **Neural pruning strategies**

Once we know the neural sensitivity for each neuron, we begin to study the role of specialists and generalists in odor classification using neural pruning. This neural pruning is based on synaptic pruning, which is the process that eliminates excessive or inappropriate synapses to form proper synaptic connections during the development of neurons [25, 179]. In our case, we "prune" neurons according to their neural sensitivity by assigning their maximum limit threshold (which will prevent these neurons from responding to stimuli). So we can observe what happens to the odor classification when only neurons with a certain sensitivity are able to transmit the odor patterns.

To analyze the relevance of specialist and generalist neurons in odor processing, we designed three pruning strategies:

- $G \rightarrow S$: For this modality, we started pruning those neurons with higher sensitivity (generalists) and ended with those with lower sensitivity (specialists).
- $S \rightarrow G$: In this case, we started pruning the specialist neurons and ended with the generalist ones.
- Random: To analyze whether the order of pruning matters, we also performed a random pruning.

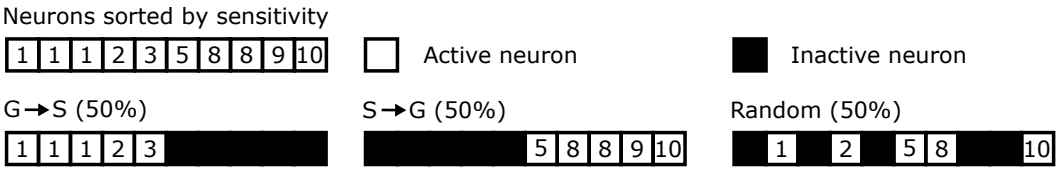


Figure 3.5: Examples of neural pruning. Given an array where we have stored the neural sensitivity of the KCs, we can sort that array according to the sensitivity (keeping the identifiers of the KCs to which these sensitivities are associated). Once the array is sorted, we can start pruning those neurons with greater sensitivity and end with those of less sensitivity ($G \rightarrow S$), we can prune in the opposite direction ($S \rightarrow G$) or prune randomly (Random).

3.2.2.3 Selection criteria of generalist and specialist neuron

Although synaptic pruning showed that specialists were more relevant than generalists for odor classification [145], we also found that the neural sensitivity distribution for KCs that achieved the minimum classification error varied according to the complexity of patterns (see section 4.2). These results suggested that the ratio of specialist and generalist neurons in KCs could be an animal adaptation to the complexity of stimuli in the natural environment. To investigate whether the ratio of specialists and generalists changes with the stimulus complexity, we needed to differentiate specialists and generalists by neural sensibility and control their ratio in the KCs layer.

Because the definition of specialists and generalists gives no indication about what neural sensitivities they should have (except that we can exclude neurons with

0 sensitivity) we started by considering as specialists those neurons that respond only to one odorant class and generalists those that respond to all of them. Once this extreme case was defined, we progressively expanded the neural sensitivities of each type of neuron to cover all of them and divided the histogram of sensitivities into two halves, as in panel (a) of Fig. 3.6. The advantage of using different combinations of sensitivities for specialists and generalists is that we can contrast their results and reinforce the conclusions derived from them.

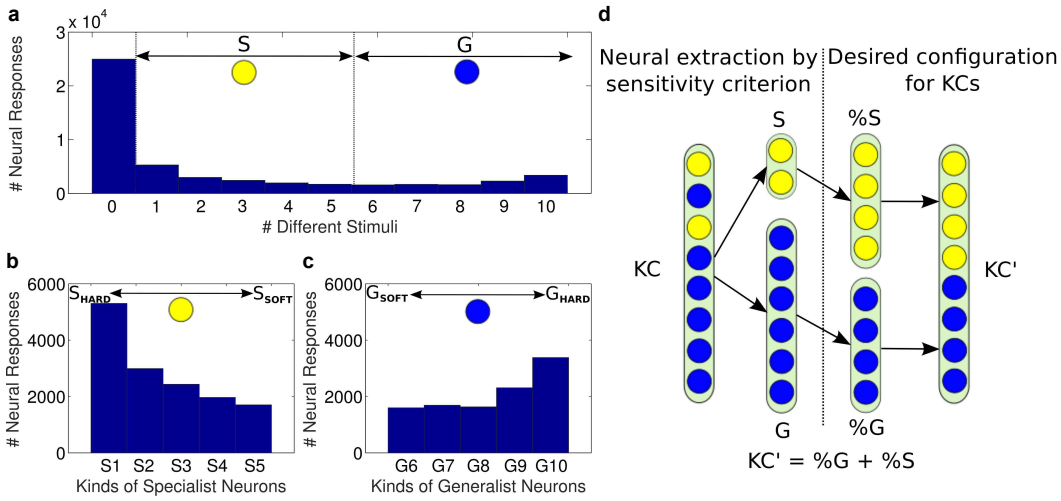


Figure 3.6: Generation of KC layer according to the selection criteria of generalist and specialist neurons. We defined specialist (S) and generalist (G) neurons by neural sensitivity, panel (a). To perform this, in the case of Gaussian patterns, we initially defined that specialist neurons had a neural sensitivity of 1 and generalist neurons had a neural sensitivity of 10. Subsequently, we started to move these boundaries (hard specialist and generalist neurons, S_{hard} and G_{hard}) to the center (soft specialist and generalist neurons, S_{soft} and G_{soft}), until we covered all neural sensitivities and specialist neurons responded from 1 to 5 different stimuli and generalist ones could do it from 6 to 10, panels (b,c). Once these boundaries had been defined, we extracted generalist and specialist neurons from KCs and we made two sets with them. Then, a new KC' layer was generated, with the same dimensions than the original one and the desired percentage of generalist and specialist neurons. In panel (d), we can see an example of an equal percentage for both types of neurons.

To try to describe the previous process more clearly, we have taken as example the Gaussian patterns and the 10 classes used for them (panel (a) of Fig. 3.6). First,

we defined specialist neurons (S) as those with neural sensitivity of 1 and generalist neurons with neural sensitivity of 10, $S1/G10$. Subsequently, we started to move the specialist and generalist boundaries to the center (panels (b,c) of Fig. 3.6), until we had specialist neurons with neural sensitivity from 1 to 5 and generalist neurons with neural sensitivity from 6 to 10, $S1-5/G6-10$.

Once specialist and generalist neurons had been defined, we made two sets of each type of neurons, as we can see in panel (d) of Fig. 3.6. Then, we created a new KC layer with the same dimensions than the original one by extracting neurons of these sets, which allowed us to control the percentages of these two types of neurons. The neural network started with all generalist neurons and we gradually replaced them by specialist neurons, in order to analyze the evolution of classification error during this process and, therefore, the relevance of this types of neurons in the pattern recognition process [146, 148, 149].

3.2.3 Gain control

Experimental studies in insects showed that the activity in the PNs remains nearly constant despite large variations of stimulus intensity [168, 207]. Therefore, a gain control mechanism [191, 201] regulating neuronal activity in the AL is likely to exist [167, 168, 235]. This mechanism prevents neurons from suffering damage or even dying by hyper-excitation [201] and can also help achieve more efficient neural code for odors, because it may tend to decrease cross-correlations between the output of different glomeruli [167, 168, 235].

In the first studies of this thesis, focused on firing thresholds and the populations of specialist and generalist neurons, we introduced this gain control mechanism in the simplest way possible, by normalizing the input patterns [144, 145, 146, 147, 148, 149]. However, we later decided to study this gain control and develop a bioinspired implementation of it [150].

3.2.3.1 Homogeneous gain control

The first computational model for the gain control mechanism that we developed only took into account the fact that the activity regulation in PNs was produced through the inhibition received by LNs (see panel (a) in Fig. 3.7).

To develop our gain control mechanism, we assumed that the odor information received by PNs and LNs is proportional [167, 168, 235], which means that, at the population level,

$$\sum_{j=1}^{N_{PN}} PN_j \approx \alpha \sum_{i=1}^{N_{LN}} LN_i \,, \tag{3.23}$$

where PN_j and LN_i are the neural activities of the j-th PN neuron and of the i-th LN neuron, respectively. Finally, α is the ratio between PN and LN population activities.

We modeled the lateral inhibition using the following non-linear relation,

$$PN_j = PN_j - PN_j\beta \left(\sum_{i=1}^{N_{LN}} LN_i - \delta \right), \quad j = 1, \ldots, N_{PN} \,, \tag{3.24}$$

where $\beta \sim 1/\sum_{i=1}^{N_{LN}} LN_i$ is the weight of the inhibitory connection from LNs to PNs, and δ is a threshold. Finally, the non-linear multiplication of $\sum_{i=1}^{N_{LN}} LN_i$ and PN_j assures that PNs with lower activities will remain with lower activities.

This first implementation of the gain control mechanism allowed us to regulate the activity of the PNs in a very similar way to the normalization of patterns (see Figs. 3.8 and 3.9). However, this method has two major problems from the biological point of view:

- First, in order to achieve the activity regulation of PNs, it was necessary that the inhibition of LNs was calibrated according to the activity of each glomerulus (symbolized by PN_j in equation 3.24 as pre-inhibitory activity) and the strength of the synaptic connection β (which varies depending on the total activity). Both terms of calibration are used to emulate, in our spatial model, the progressive calibration of the activity in PNs through their spatio-temporal dynamics, but we have not found any biological evidence to support them.

- Second, biological recordings showed that when the activity of ORNs is low, the activity of PNs is proportional to it. But, when the activity of ORNs reaches a certain activity level, the activity of PNs saturates. However, in our mechanism, the activity of PNs remain stable independently of the activity

of the ORNs. This behavior, similar to the normalization of patterns, is more strict than the one observed in AL.

3.2.3.2 Heterogeneous gain control

In order to develop a gain control mechanism closer to the biological one, we decided to take into account the existence of two types of LNs: homoLNs and heteroLNs. While homoLNs innervate most if not all glomeruli uniformly, heteroLNs innervate only a few of them [3, 38, 46, 65, 133, 193, 212]. The introduction of these neurons transforms the homogeneous inhibition of our previous gain control mechanism in a heterogeneous inhibition that can be the real responsible for the gain control mechanism, as suggested in [137].

To implement this heterogeneous gain control mechanism (see panel (b) in Fig. 3.7), we took into account that the proportion of the LNs is approximately one third of the PNs (830 PNs [116] and 300 LNs in locust [3], $150-200$ PNs [205] among the 250 AL neurons in *Drosophila* [202]) and made the following assumptions: (i) all PNs are uniglomerular (as it happens in the majority of PNs of the *Drosophila*), (ii) homoLNs are connected to all the glomeruli and use the activity received by one of them to inhibit the rest, ensuring that the odor information of a glomerulus is received by at least one group of homoLNs, and (iii) heteroLNs are connected randomly to a few glomeruli and use their average activity to inhibit them. This gain control mechanism can be expressed mathematically as follows:

$$PN_p^{GL_i} = PN_p^{GL_i} - \alpha Inhib_{homo} - \beta Inhib_{hetero}, \;\; p = 1, \ldots, N_{PN}\,, \tag{3.25}$$

$$Inhib_{homo} = \left(\sum_{j \neq i}^{N_{GL}} \sum_{k=1}^{N_{homoLN}{}^{GL_j}} homoLN_k^{GL_j} \right)\,, \tag{3.26}$$

$$Inhib_{hetero} = \sum_{l=1}^{N_{heteroLN}{}^{GL_i}} heteroLN_l^{GL_i}\,, \tag{3.27}$$

where $PN_p^{GL_i}$ is the p-th PN that innervates in the i-th glomerulus, $homoLN_k^{GL_j}$ is the k-th homoLN that receives the odor information of the j-th glomerulus and inhibits the rest (among them the i-th glomerulus) and $heteroLN_l^{GL_i}$ is the l-th

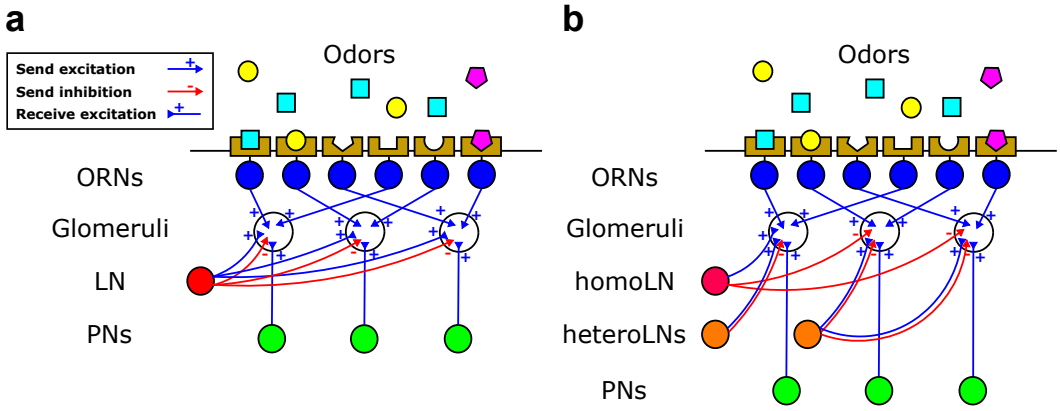


Figure 3.7: Schematics of the gain control mechanisms implemented. Panel (a) shows the schematic of the gain control described in equation 3.24. In this gain control mechanism, LNs receive excitation from all the glomeruli and subsequently inhibit them depending on the total activity received. This inhibition is calibrated according to the activity of each glomerulus (symbolized by the PN activity) and the strength of the synaptic connection (which varies depending on the total activity). Panel (b) displays the schematic of the gain control described in equation 3.25. In this gain control mechanism, there are two types of LNs: homoLNs and heteroLNs. Although homoLNs innervate all glomeruli, we suggest that these connections are not identical and homoLNs receive excitation by a single glomerulus and inhibits the rest of them as a function of its activity. This type of connections in homoLNs allows the most active glomeruli dominate over the less active ones, something that (after several tests) we saw that it was necessary for the proper performance of the gain control mechanism. On the other hand, heteroLNs receive excitation from a few glomeruli and subsequently inhibit them depending on their averaged activity.

heteroLN that innervates in the i-th glomerulus (see panel (b) in Fig. 3.7). On the other hand, N_{GL} represents the number of glomeruli, $N_{homoLN^{GL_j}}$ the number of homoLNs that receive the odor information of the j-th glomerulus and $N_{heteroLN^{GL_i}}$ the number of heteroLNs that innervate in the i-th glomerulus. Finally, the coefficients α and β allow regulating the inhibition of the two types of LNs ($Inhib_{homo}$ and $Inhib_{hetero}$) for achieving a suitable gain control.

To calculate the values of α and β we used 500 Gaussian patterns divided between 5 concentrations $(5, 25, 50, 75, 100)$ and 10 pattern classes similar to those shown above in Fig. 3.12. The dimension of these patterns is based on the *Dro-*

sophila: 180 PNs and 80 LNs [157, 205]. From this value for LNs we assume 40 homoLNs and 40 heteroLNs, where heteroLNs can be randomly connected to 1, 5 or 10 glomeruli (with the aim of simulating different degrees of innervation). The results of this calculation were $\alpha = 1$ and $\beta = 0.04$.

3.2.3.3 PNs activity obtained by the different gain control mechanisms

Once we made these two models of gain control mechanism, we wanted to observe their effects on the activity of the PNs. Hence, in Fig. 3.8 we can observe the activity generated in these neurons by patterns with different concentration levels in case of not using any gain control mechanism or one of the three mentioned above. The resulting activity in PNs for these four scenarios are the following:

- When there is no gain control, the activity of PNs increases with the concentration level, each time with less intensity until it saturates when the odor concentration makes that all PNs fire with a similar intensity. For this extreme case, the Gaussian distributions used as input pattern will look as uniform distributions, which represents the unreal situation where all glomeruli are highly active. Because of this, we do not take into account this extreme case for our experiments and Figs. 3.8 and 3.9.

- When we use a homogeneous gain control mechanism as the normalization of patterns or our first model, Eq.3.24, PNs always have the same activity independently of the concentration (see Norm and Homo lines in Fig. 3.8). A behavior different from that observed in AL, where the activity of PNs grows for low values of concentration until it reaches a certain concentration level for which this activity becomes more or less stable [167, 168, 235].

- This biological behavior is only achieved when we use our heterogeneous gain control mechanism, Eq.3.25, as can be observed in Hetero lines of Fig. 3.8.

How these three gain control mechanisms described affect the activity of PNs can be observed in Fig. 3.9. In the first column of this figure, we present an example of a pattern for each of 5 different concentration levels $(5, 25, 50, 75, 100)$ without applying any gain control mechanism for 3 different noise levels (low,

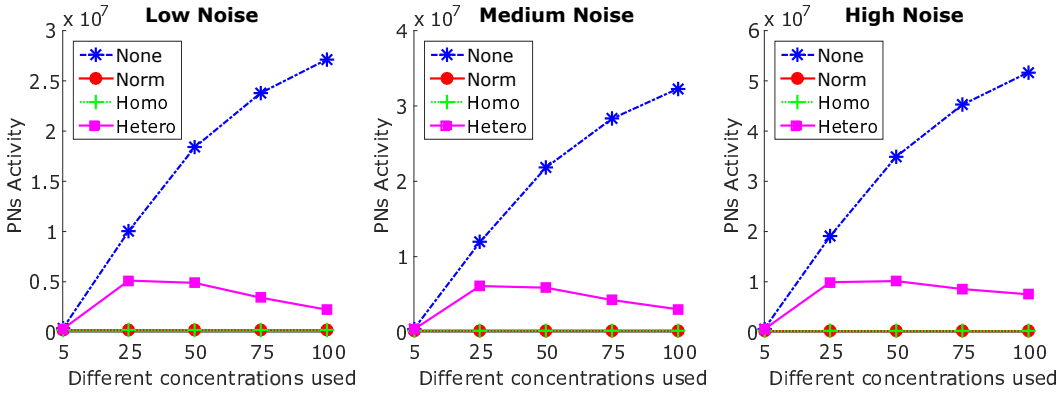


Figure 3.8: Activity in PNs according to odor concentration and gain control mechanism used. When there is no gain control, the activity of PNs increases with the concentration level, each time with less intensity until it saturates when the odor concentration makes that all PNs fire with a similar intensity. When we use the gain control based on the combination of heteroLNs and homoLNs (Hetero), we obtain a behavior similar to the one observed biologically [167, 168, 235], where the activity grows for low values of concentration until it reaches a certain concentration level for which the activity becomes more or less stable. This biological behavior does not occur for the normalization and our homogeneous gain control mechanism (Homo), in which the PNs always have the same activity independently of the concentration.

medium and high). We can see that the effects on Gaussian patterns produced by the normalization and our homogeneous gain control mechanism (Homo) (second and third column) are similar and do not correspond with the biological behavior observed in AL. Finally, the fourth column shows the activity of PNs when we use our heterogeneous gain control mechanism (Hetero), which allows distinguishing the activity focus of PNs more clearly.

3.3 Input patterns

To understand how the insect olfactory system processes the odorants and find out what properties of their neural network are involved in their classification, we needed different sets of patterns to simulate the odor coding in AL and to help us achieve our different objectives. The first set of patterns used was developed by us in order to make a first approach to the insect olfactory system and the computa-

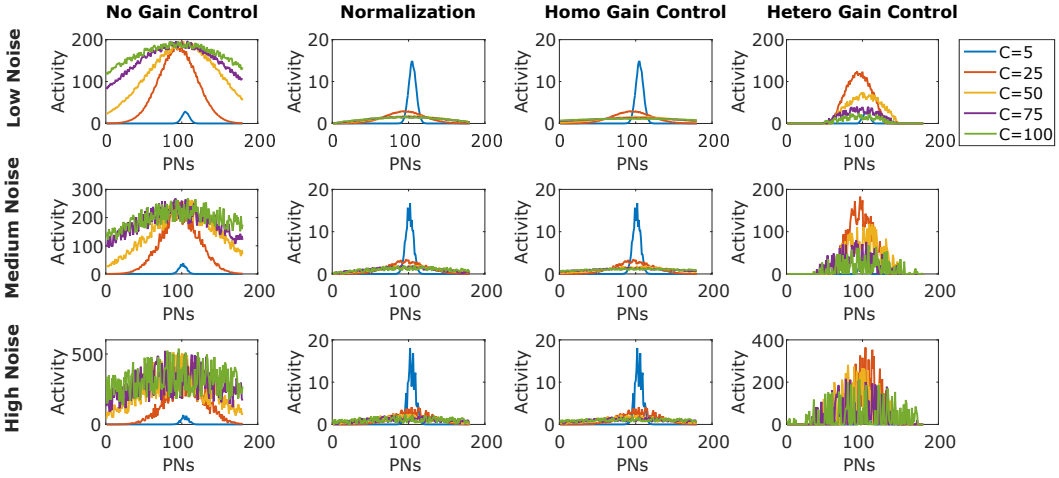


Figure 3.9: Gain control effects on the activity of PNs. The first column shows an example of a pattern for each of 5 different concentration levels without applying any gain control mechanism, displaying these examples for 3 different noise levels in each row. In the second and third column (Normalization and Homo Gain Control), we can observe that our homogeneous gain control mechanism (Eq.3.24) obtained a similar activity in PNs to the one obtained for normalization. Finally, the fourth column (Hetero Gain Control) displays the activity of PNs when we use our heterogeneous gain control mechanism (Eq.3.25), which allows distinguishing the activity focus of PNs more clearly.

tional model that simulates it. These patterns of low complexity, which we called primary patterns, are divided into two types depending on whether or not interclass overlap exists: orthogonals and characters. These patterns, used in [144, 146], were later replaced, for the rest of the thesis, by Gaussian patterns, with the aim of simulating the spatial activity patterns observed in the AL glomeruli [53, 208], giving us great control of the overlap degree between different pattern classes. This overlap control allows us to generate patterns with different complexity levels and, therefore, to study the odor processing in the insect olfactory system under different difficulty degrees. In addition to these patterns developed by us, we use two external databases: the MNIST digits [115] and chemical data obtained through electronic noses [184, 225]. The objective of using these external datasets is to study how the computational model used solves a typical problem of pattern recognition, such as the MNIST digits, and classifies data of chemical compounds,

which make up the odorants.

3.3.1 Primary patterns

As we mentioned above, to make a first approach to the computational model and the odor classification problem, we used two simple kinds of patterns: orthogonals and characters (see Fig.3.10). The coding of both patterns is binary and, therefore, the PNs they represent are considered active (1) or inactive (0). Orthogonal patterns are those that do not have interclass overlap, while character patterns (which after preliminary tests became only digits) have this kind of overlap but not interclass overlap, unlike MNIST digits, making them a good first step to later classify MNIST digits. However, this lack of interclass overlap in orthogonal patterns and the intraclass one in characters is modified when we introduce some noise in them, which adds some difficulty to their classification. These patterns were used in [144, 146], whose results are shown in section 4.1, being later replaced by Gaussian patterns that allow us to control the overlap degree of patterns more gradually and not only by introducing noise.

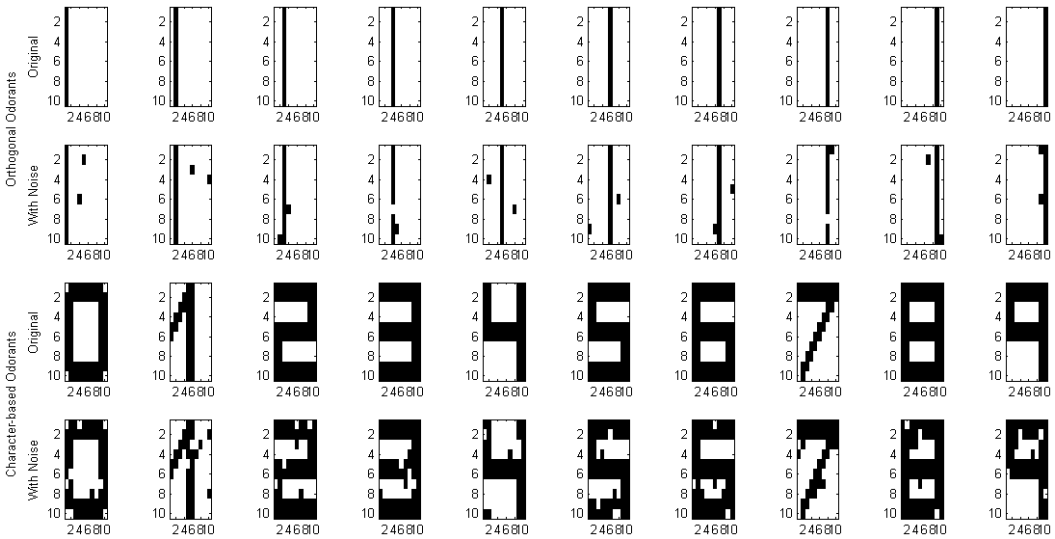


Figure 3.10: Examples of orthogonal and character patterns. Odd rows represent these two types of patterns without noise, while even rows show how noise has been included in them (these modified patterns are the ones we will use in our experiments). Colors: black (active neuron), white (inactive neuron).

3.3.2 Gaussian patterns

Different odorants activate different regions of AL [187]. Different levels of odor concentration can expand the number of AL neurons that are activated for a certain odorant [187, 190, 208]. A higher concentration of odor leads to better detection by the olfactory system [238]. However, we can assume that if the activity region for an odor pattern grows without control, the overlap with the regions used by other odorants will also increase.

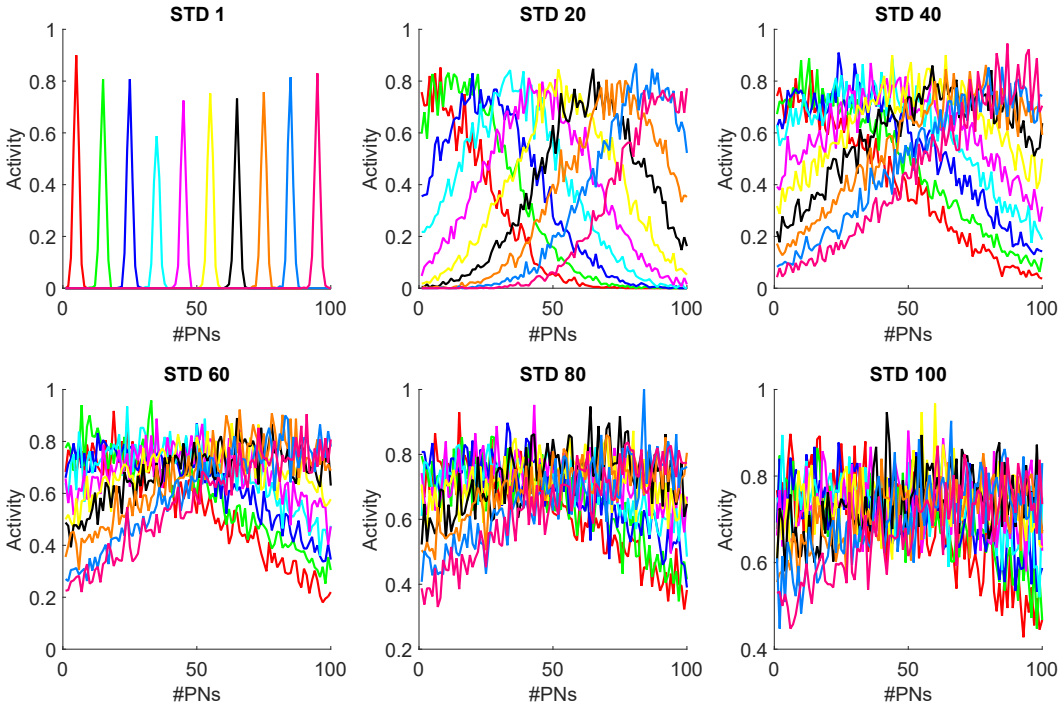


Figure 3.11: Gaussian patterns. In these panels, we can see an example of a pattern for 10 pattern classes and 6 standard deviations (STD). The X-axis of the graphs represents the PNs, while the Y-axis represents the activity level (firing rate) of these neurons. When odorants are encoded by a fewer number of neurons (low STD) the firing rate of these neurons increases. On the other hand, when a greater number of neurons encode the odor patterns, the firing rate is distributed among them, increasing their STD and noise.

Considering this, we used Gaussian patterns centered at different input neurons depending on the class to which they belong (Fig.3.11, where we can see a pattern example for each of their classes and configurations). These patterns represent

the firing rate of PNs caused by different odorants, which can involve a variable number of neurons. This variable number of PNs stimulated by the odorants is controlled by standard deviation (STD), which implies different overlapping degrees and different noise degrees.

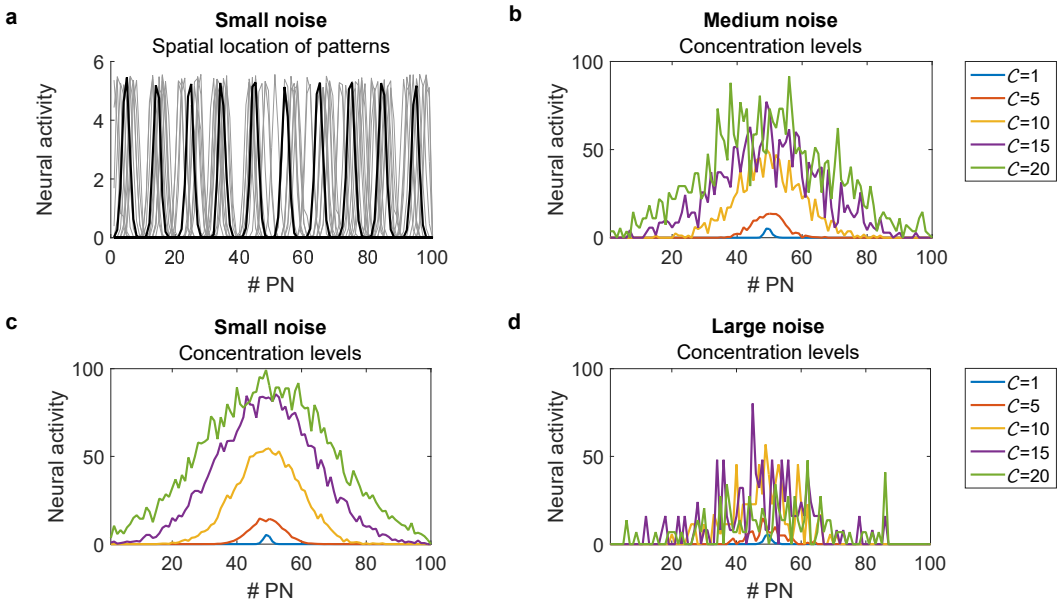


Figure 3.12: Gaussian patterns for studying the gain control mechanism. **(a)** Different odor identities are modeled as Gaussians centered at different PN neurons ($\mathcal{C} = 1$). To generate more realistic stimulation, we added noise. **(b)** Neural activity variability as a function of concentration. **(c)** Activity in the PN neurons elicited by an odor with small level of noise. **(d)** Activity in the PN neurons for a larger level of noise.

In order to use the Gaussian patterns to study the gain control mechanism present in AL, we decided to develop a new method to create these patterns by using the following equation:

$$PN_j = \left[\mathcal{A}_0 + \frac{\mathcal{A}}{1 + \left(\frac{K}{\mathcal{C}}\right)^Q}\right] \exp\left(-\frac{(j - j_k)^2}{2\mathcal{C}}\right), \tag{3.28}$$

where $\mathcal{C}$ is the concentration, j_k is the center of the j-th odor identity, $\mathcal{A}_0 = 0.05$ is the residual activity, $\mathcal{A} = 0.95$ is a parameter that determines the maximum activity, $K = 10$ defined the concentration at which half of the maximum level was

achieved, and $Q = 3.2$ defined the slope of the resulting sigmoid. This follows a common psychophysical response of neural populations to different intensities [152].

We used on total five different levels of concentration (Fig.3.12). Each of these concentrations scales the pattern by a sigmoidal function, which changes the variance of the Gaussian because neighboring PN neurons also become active (in agreement with experiments [187, 190]). Thus, higher concentrations increase the overlap between the activities generated by two different odors. Finally, to simulate the noise experienced in real biological systems, we added noise to the Gaussian center (Fig.3.12a) and to the stimulus itself (Fig.3.12b-d).

The versatility of these patterns due to the possibility of controlling their complexity makes them the most used in this thesis [145, 148, 149, 150]. These patterns have been used to study specialist and generalist neurons according to the input complexity and the gain control mechanism in AL (section 4.2-4.5).

3.3.3 MNIST patterns

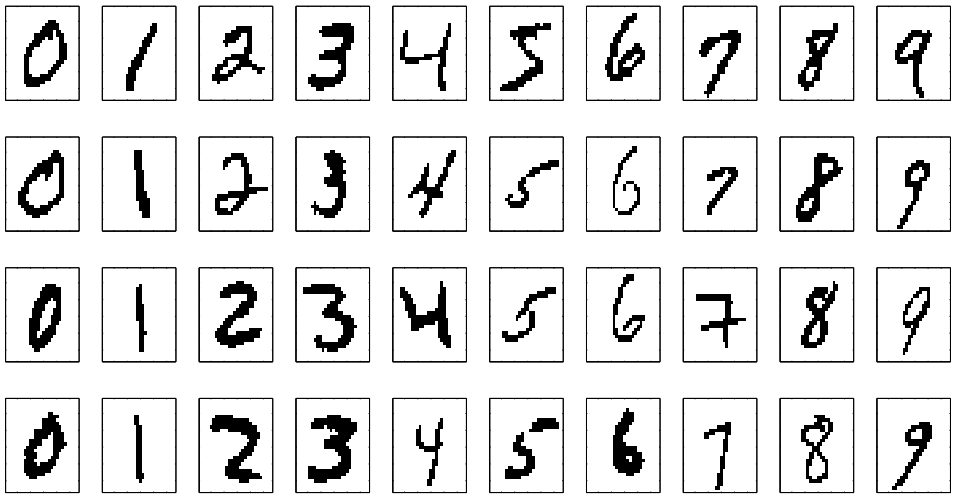


Figure 3.13: Example of MNIST digit patterns. In each column, we can observe the 10 pattern classes and in each row an example of these patterns. The active neurons, as happened in the primary patterns, correspond to the black pixels, while the inactive ones to the white ones.

The MNIST database of handwritten digits [115] is a well known dataset and widely used for analyzing pattern recognition algorithms. In fact, this dataset was used in other study for the same computational model [89] and that allows us to compare our results with them (in [147] we compare the previous results obtained in [145], while in [148] the comparison is made in the same study). The MNIST digits have dimensions of 28×28 pixels (Fig.3.13), which we converted into an array of 784 elements (PNs) by merging each row with the next.

3.3.4 Electronic noses data

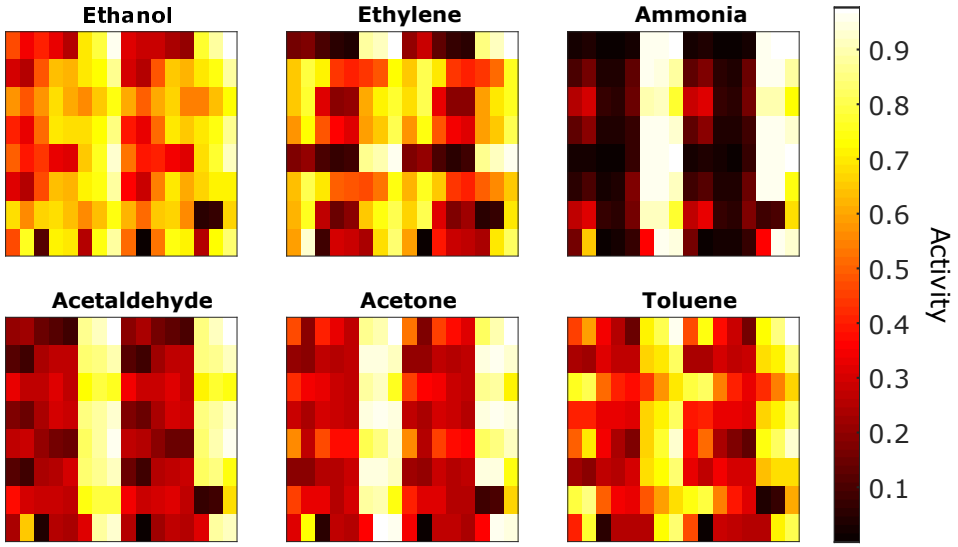


Figure 3.14: Electronic noses patterns. Representation of the 6 pattern classes through the normalized mean activity of their 128-dimensional feature vector.

The second external data set used to compare the results obtained by primary patterns [144] and Gaussian patterns [148] are chemical compound data from electronic noses, because it gives us an odor coding that can offer certain similarity with the one made by the olfactory system. The data used belongs to BioCircuits Institute [184, 225], University of California in San Diego, that contains 13,910 measurements from 16 chemical sensors exposed to 6 gases at different concentration levels. Each sensor extracted 8 features from gases, resulting in a 128-dimensional

feature vector (8 features x 16 sensors). The gaseous substances used are Ammonia, Acetaldehyde, Acetone, Ethylene, Ethanol, and Toluene (Fig.3.14), dosed at a wide variety of concentration levels in the intervals (50,1000), (5,500), (12,1000), (10,300), (10,600), and (10,100) ppmv, respectively. The data set contains patterns corresponding to three years of measurements divided into 10 batches, which implies variability in the chemical detection between batches due to deterioration of the sensors caused by aging and contamination. Because of this, we decided to choose a single batch, batch 10, that contains the largest number of samples for each gaseous substance (see Table 3.1). However, since not all the chemicals were measured at the same concentrations, we chose the concentration that included the largest number of samples, 50 ppmv (100 samples for each class). The lack of enough samples for different concentrations was the fundamental reason why we use the Gaussian pattern when we want to simulate different concentration conditions. We hope to solve this problem in the future by searching a new chemical data set or the reduction of classes in this one, in order to complete our study about the gain control mechanism in AL.

Batch	1	2	3	4	5	6	7	8	9	10
Ethanol	83	100	216	12	20	110	360	40	100	**600**
Ethylene	30	109	240	30	46	29	744	33	75	**600**
Ammonia	70	532	275	12	63	606	630	143	78	**600**
Acetaldehyde	98	334	490	43	40	574	662	30	55	**600**
Acetone	90	164	365	64	28	514	649	30	61	**600**
Toluene	74	5	0	0	0	467	568	18	101	**600**

Table 3.1: Chemical samples for each batch. For all these batches, batch 10 has the largest number of samples and is the most balanced in terms of samples per class.

CHAPTER

4

Results

In order to understand how the changeful information of odorants (different mixtures, concentrations, environment conditions, etc.) [54] is processed by the olfactory system, we focused on the simple structural organization that this system has in the insects [5, 43, 58, 59, 78, 95, 140, 209]. Using a single hidden layer neural network that retains the main parts of the insect olfactory system involved in odor classification [89], we have analyzed a property barely studied and usually not applied in biological models: neural heterogeneity. In particular, we aim to analyze three types of heterogeneity observed in the insect olfactory system:

- The variability of neural thresholds in KCs [175].
- The diversity observed in the neural sensitivity of KCs that divides its neural population between specialist and generalist neurons [28, 175, 182, 183, 247].
- The inhibitory heterogeneity that produces the gain control mechanism observed in AL [65, 193].

In section 1 of this chapter, we show the results of our study of neural thresholds. In this study, as we mentioned in the previous chapter, we compare the classification errors obtained by using homogeneous thresholds (a single threshold value

for all neurons) and heterogeneous thresholds (a different value for each neuron). The selection of values for homogeneous and heterogeneous thresholds was made by two different methods: exhaustive search and gradient descent algorithm. The results obtained by the exhaustive search are detailed in section 4.1.1 and published in [144, 146], while in section 4.1.2 we discuss preliminary results for the gradient descent algorithm (using the equation 3.12). For this last method, we are working on a new algorithm based on the equation 3.22. Early results are discussed in [123, 124].

Sections 2, 3 and 4 present the study on specialist and generalist in KCs and, therefore, the heterogeneity of neural sensitivity. In section 2, we verified by neural pruning the fundamental role of specialist neurons in odor discrimination. The results, published in [145], suggested a variable relevance of specialists according to the complexity of patterns. Therefore, in section 3, we estimated the ratio of specialists and generalists necessary for the system to reach the minimum classification error. This study showed that there is a relationship between the ratio of these neurons and the complexity of the patterns, with some configurations for which the presence of generalist neurons was also necessary [146, 148]. Finally, in section 4, we try to determine whether the balancing process performed in the previous section altered the randomness of the connections between PNs and KCs, in order to know if this process of adaptation to the complexity of the stimuli is possible in biological systems. The results showed that this balance does not vary the random connectivity, because of neural sensitivity is not related to the number of connections that KCs receive [149].

Finally, section 5 shows the preliminary results for the computational model of a heterogeneous gain control mechanism based on the populations of inhibitory neurons: homoLNs and heteroLNs. This configuration for gain control seems to lead to a better odor classification with respect to homogeneous gain control mechanisms, such as the one developed by us in [149]. However, both types of gain control allow the olfactory system to noticeably improve its classification capability.

4.1 Neural threshold heterogeneity improves odorant classification

To study the neuronal diversity of firing thresholds in odor processing, we compare the classification error obtained by using homogeneous and heterogeneous thresholds. To calculate these threshold values we proposed two different methods: exhaustive search and gradient descent algorithm.

4.1.1 Threshold selection by exhaustive search

For an initial approach to study the heterogeneity of neural thresholds [144], we used only the primary patterns (see section 3.3.1). We used 4 sets of patterns, 2 for each type of primary pattern (orthogonal and character) consisted of 15 and 30 patterns (5 and 10 pattern classes, 3 patterns for each class) with a 20% of noise. For each set, we ran 10 simulations using 3-cross-validation. These simulations were done for different probabilities for PNs-KCs connections (p_c). We used different connection probabilities for p_c $(0.1, 0.3, 0.5)$ based on studies that confirm that these are feasible values for it [62, 101] and a probability $p_w = 0.5$ because W matrix is subjected to Hebbian learning (see section 3.1.4.1). The probabilities of reinforcement and weakening of weights, for this learning, are $p_+ = 0.2$ and $p_- = 0.1$ respectively (see equation 3.3).

The averaged results for these set of patterns (Fig. 4.1) show that heterogeneous thresholds achieve lower classification errors. Furthermore, we observe that lower connection probabilities p_c are also related to lower errors.

Once we observed the relationship between classification error and heterogeneous thresholds for different connection probabilities, we analyzed the relationship between classification error and spike rate (neural response level) for a particular case, $p_c = 0.1$ and 15 patterns (spike rate can be observed in the dotted line in Fig 4.2).

These results show that minimum classification error is related to a low spike rate. This happens because, if the spike rate is high, the overlap between patterns of different classes increases, which hinders correct classification. However, if the

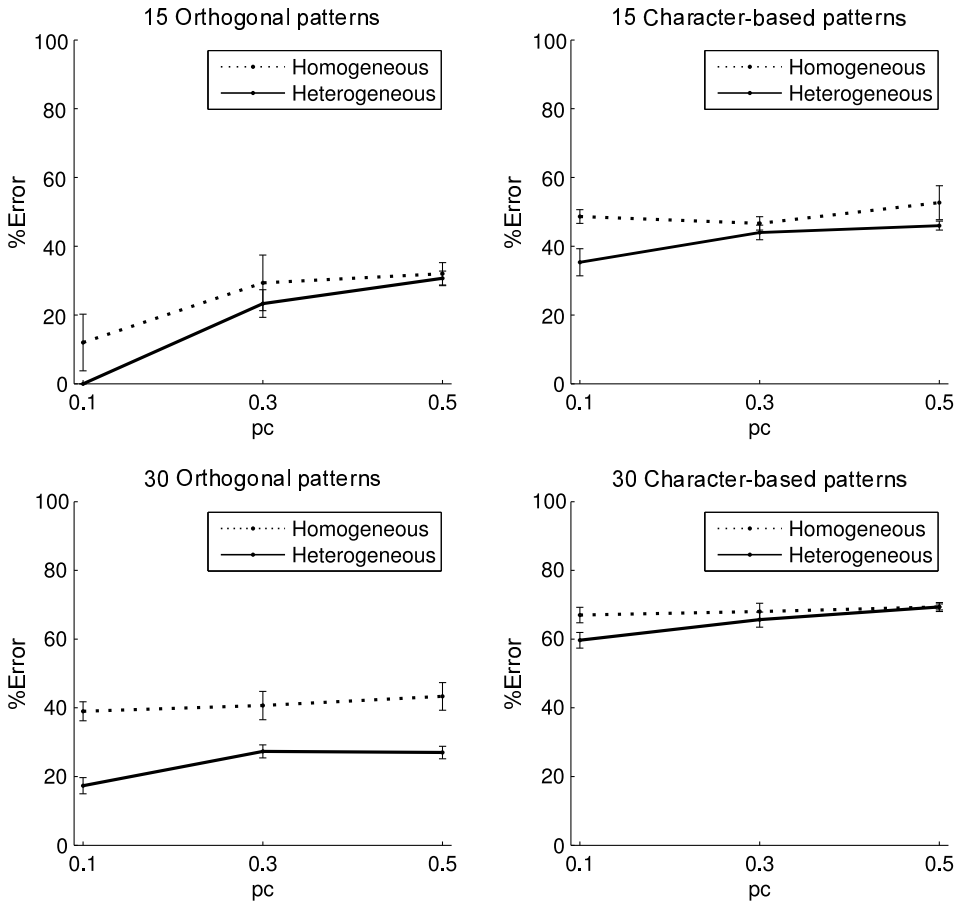


Figure 4.1: Comparison of error rate for homogeneous and heterogeneous thresholds We compared the results obtained by homogeneous and heterogeneous thresholds for orthogonal (first column) and character patterns (second columns) using 15 and 30 patterns (top and bottom panels) and for different connection probabilities for PNs-KCs connections, p_c. Sample means with 95% confidence intervals of standard errors (SE).

spike rate drops excessively, the odor information that arrives at the output neurons is insufficient for its classification.

In order to generalize these preliminary results, we increased the set of patterns (adding the electronic noses data), the number of patterns used by each set (100) and the number of simulations performed (100 with 5-cross-validation) for each data set and connection probability p_c [146].

The introduction of electronic nose data (shown in section 3.3.4) was a chal-

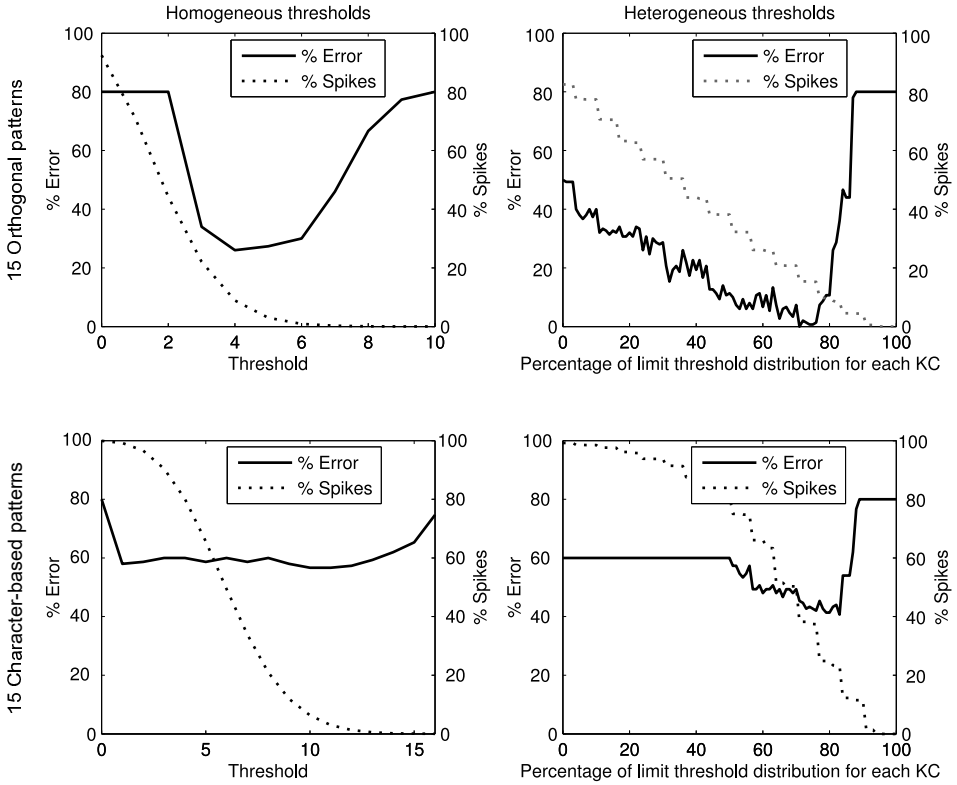


Figure 4.2: Relationship between classification error and spike rate. Results for 15 orthogonal and 15 character-based odorants with noise and connectivity probability $p_c = 0.1$. In the case of homogeneous thresholds, the X axis represents the value taken for the hidden layer thresholds. For heterogeneous thresholds this axis represents the percentage of odors for which the KCs do not respond (each KC has a different threshold value according to its limit threshold distribution).

lenge for the algorithm of heterogeneous threshold selection used in the previous case (see algorithm 2 in section B.1.2.1). To analyze why heterogeneous thresholds obtained better classification results than homogeneous ones for primary patterns but not for electronic noses data (for which both thresholds obtained similar results), we investigated the limit threshold distributions for each neuron. This analysis showed that there were neurons with little variance between their limit threshold values and others with a large variance, whose distribution tended to be bimodal (see Fig. 3.3). This fact, consistent with the generalist and specialist neurons observed in KCs, led us to develop a new algorithm for heterogeneous

threshold selection that could use these types of neurons in order to be more efficient. The use of this new algorithm (see algorithm 3 in section B.1.2.2) again showed an improvement in odor classification for heterogeneous thresholds and low p_c probabilities, like $p_c = 0.1$ (Fig. 4.3).

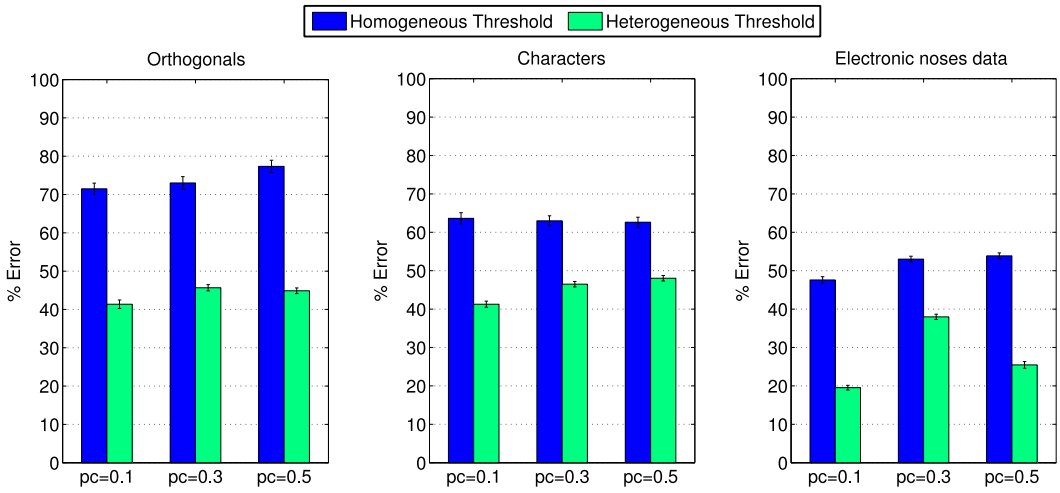


Figure 4.3: Comparison between homogeneous and heterogeneous thresholds for different sets of patterns and connection probabilities between PNs and KCs (p_c). Sample means with 95% confidence intervals of standard errors (error bar calculated over 100 simulations).

Since the electronic noses data has a smaller number of classes (5) with respect to the orthogonal and character patterns (10) and a greater number of patterns per class (20 instead of 10), we obtained the best classification error for this data set. Instead, orthogonal patterns, theoretically simpler, are those that achieve a higher classification error, since the noise introduced in these patterns is maximized due to the small number of neurons that encode their information.

If we analyze again the relationship between classification error and spike rate (Fig. 4.4), for $p_c = 0.1$, and compare it with the previous results (Fig. 4.2), we can observe a similarity between them. Both figures show that minimum classification error is related to a low spike rate, which is consistent with the sparse activity documented in the KC layer [175, 214]. Moreover, we can see the effects of using the new algorithm for heterogeneous threshold selection in the spike rate for orthogonal and character patterns of both figures. Because the new algorithm has

different specialization coefficients for each neuron (see the explanation of the algorithm in section 3.2.1.1), the spike rates for these patterns now start at 20% when their values were higher than 80% using the previous algorithm.

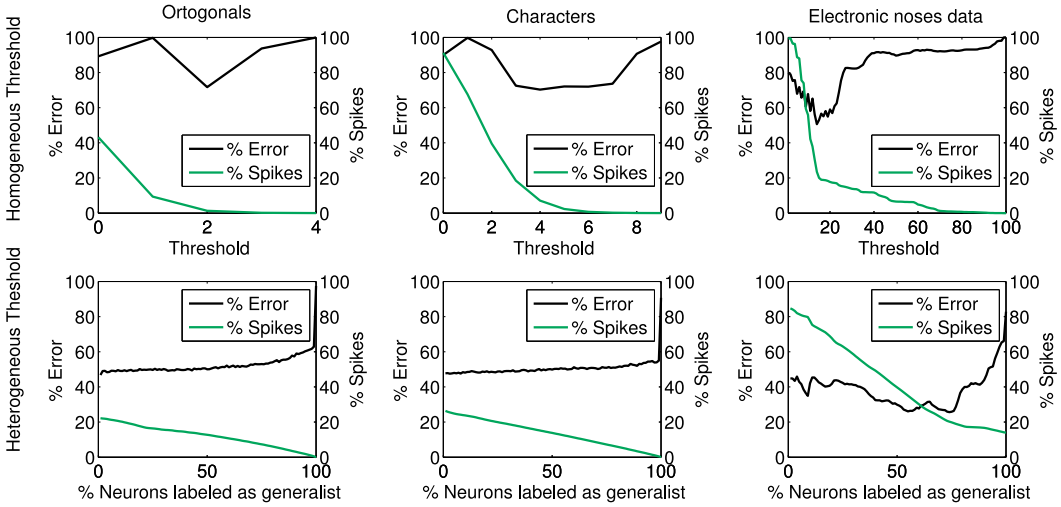


Figure 4.4: Classification error and maximum spike rate for different kinds of patterns ($p_c = 0.1$). These results are the average of 100 simulations for different C and W weights. These results show, for all pattern classes, that classification error is always lower for heterogeneous thresholds. We calculate the heterogeneous thresholds differently according to whether they are specialist or generalist neurons. In the case of generalist neurons, we assign them the maximum limit threshold (which is equivalent to prune these neurons). In the case of specialist neurons we consider their distribution as bimodal and try to find their minimum turning point.

4.1.2 Threshold selection by gradient descent

The algorithms implemented for the heterogeneous threshold selection by brute-force search (see sections B.1.2.1 and B.1.2.2) were based on approximations that used percentages to restrict the activity of KCs to certain odorants or different calculation methods for specialist and generalist neurons (see section 3.2.1.1). These approximations try to solve the problem of the high cost of checking all possible combinations of threshold values. However, the resulting methods did not allow us to calculate exactly the optimal values for homogeneous and heterogeneous

thresholds. In order to have a method that would allow it, we developed a gradient descent method, which is described in section 3.2.1.2.

To compare the heterogeneous thresholds with the homogeneous ones through their classification error, we used the same patterns used in [146] (electronic nose data and primary patterns), whose errors for the previous threshold selection method were shown in Fig. 4.3. These data sets had 5 pattern classes and 20 patterns for each class (100 patterns in total). Furthermore, the connection probability used in C matrix was $p_c = 0.1$ because of its good results shown in the last section.

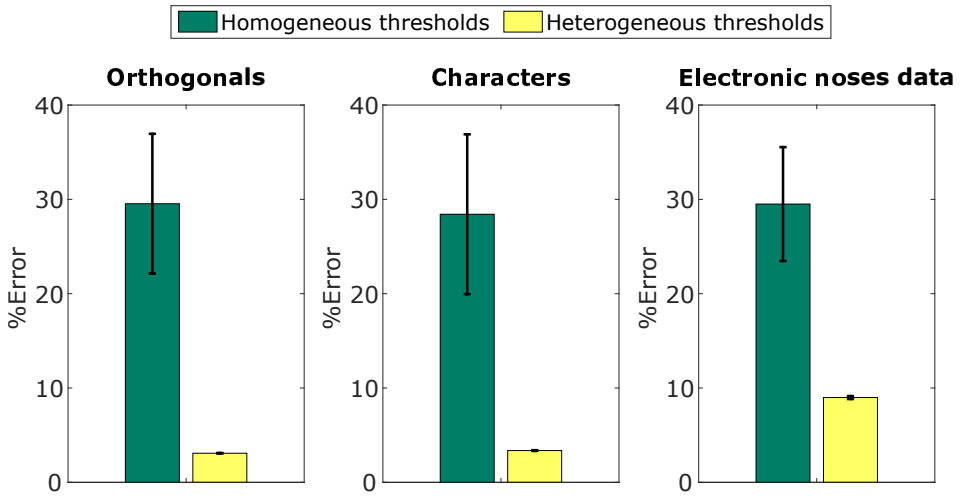


Figure 4.5: Comparison between homogeneous and heterogeneous thresholds using a gradient descent algorithm. This figure shows the results averaged for 10 simulations using 5-cross-validation for the patterns used in [146] (electronic nose data and primary patterns (orthogonals and characters)). These data sets had 5 classes and 20 patterns per class (100 in total). The classification results for the heterogeneous thresholds show an improvement of $\sim$ 24% over the results obtained for the homogeneous ones. However, the gradient descent algorithm (equation 3.22) has difficulties to correctly classify other sets of patterns such as MNIST digits (shown in section 3.3.3). Because of this, we are currently using a simplified version of this algorithm [123, 124].

These preliminary results were consistent with those obtained for the exhaustive search, showing that heterogeneous thresholds achieve better classification results than homogeneous ones. However, as mentioned in section 3.2.1.2, equation 3.12 had several problems. A high computational cost, difficulty in escaping

from local minima and difficulty in balancing the different learning terms. These problems made it difficult to work with different data sets, so we decided to simplify it 3.22. The refinement of this new equation is being accomplished by Jessica Lopez-Hazas, whose preliminary results can be found in [123, 124].

4.2 Relevant role of specialist neurons for pattern recognition

As shown in 4.1, controlling the populations of specialist and generalist neurons (observed in the KCs) was useful for the heterogeneous threshold selection by exhaustive search (see Fig. 4.4). This positive relationship between odor classification and these types of neurons led us to investigate them, as their role in the olfactory systems is still unclear [28, 102]. The common hypothesis is that specialist neurons (which respond selectively to stimuli) are crucial for discrimination, while generalist neurons (which code multiple stimuli) play a key role in extracting and discovering common features [236]. Thus, in order to check this hypothesis and observe the relevance of the specialists in odorant discrimination, we developed a method based on neural pruning for studying these types of neurons (see section 3.2.2.2).

This method progressively "pruned" the hidden layer neurons (KCs) by assigning them their maximum limit threshold, for which they can never be activated. This pruning was done in three different ways using the concept of neural sensitivity [182, 183] (the number of different stimuli to which a neuron responds, see section 3.2.2.1):

- $S \rightarrow G$: initially those neurons with a lower neural sensitivity (specialists) are pruned, being those with higher neural sensitivity (generalists) the last pruned.
- $G \rightarrow S$: we started pruning neurons with high neural sensitivity and ended with those with low sensitivity.
- Random: we perform the pruning of neurons randomly to analyze if the order of pruning matters.

We used these different strategies of neural pruning on 15 sets of Gaussian patterns (where we used 5 overlap degrees between patterns, using their standard deviation (STD), and 3 noise levels). Each set has 5 classes of Gaussian patterns with 20 patterns per each class (100 patterns in total) and it was ran 10 times for each configuration of p_c (0.1, 0.3, 0.5). The connection probability for W (p_w) matrix was 0.5 and the probabilities for the supervised Hebbian learning that modifies its weights were $p_+ = 0.2$ and $p_- = 0.1$ (see section 3.1.4.1).

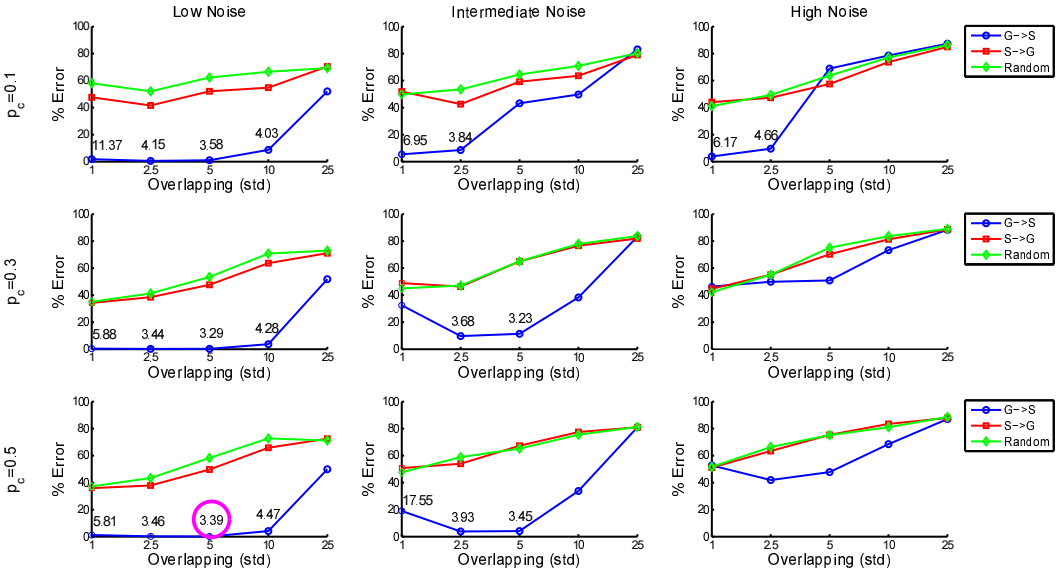


Figure 4.6: Classification error for different neural pruning and connection probabilities. Comparison between different neural pruning strategies for different overlap degrees in the Gaussian patterns (std) and connection probabilities between PNs and KCs (p_c). The values shown in some points of the error values, represent the firing rates (%) for KC neurons where the error is minimized. For example, for low noise, connection probability $p_c = 0.5$ and overlapping degree $std = 5$, the firing rate in KCs is 3.39% (see magenta circle). The $G \rightarrow S$ pruning achieves the best classification error in all cases. These results also show that the model is more robust to noise for low connectivity probabilities.

The averaged results for these simulations (Fig. 4.6) showed that the best neural pruning was from the most generalist neurons (high neural sensitivity) to the most specialist ones (low neural sensitivity), $G \rightarrow S$. This was the unique pruning strategy that minimized the classification error, confirming the great relevance of

specialist neurons for pattern recognition.

Furthermore, the firing rates (%), showed in Fig. 4.6 for those points where the error was minimized, indicate that minimum classification error was achieved when there was sparse activity (see the 3.39% of activity in KCs marked with a magenta circle in Fig. 4.6, for $p_c = 0.5$ and $std = 5$). These results are consistent with the sparseness condition observed in these KCs [175].

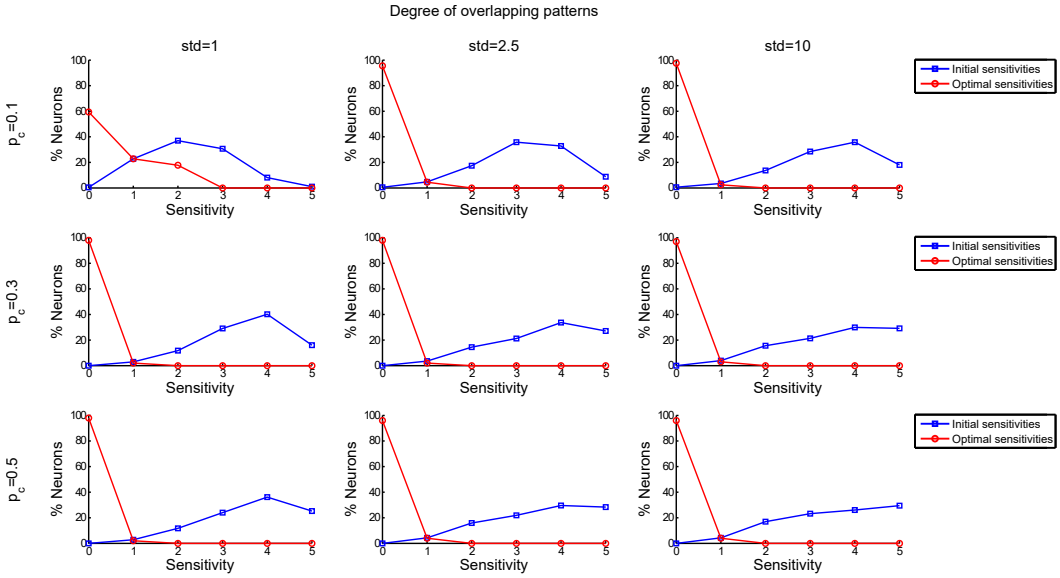


Figure 4.7: Sensitivity evolution for different configurations of $G \rightarrow S$ pruning. Relationship between initial and optimal sensitivity for $G \rightarrow S$ pruning, low noise and different conditions of connection probabilities (p_c) and overlapping degrees (std). These sensitivity curves correspond to some values of the first column in Fig. 4.6. When the overlapping degree or the connection probability increases, the initial percentage of neurons with high neural sensitivity (generalists) also rises. However, the optimal neural sensitivities for KCs, for all configurations shown, are always very low. For the majority of cases, except for $p_c = 0.1$ and $std = 1$, we only have a few neurons with sensitivity 1 and a large majority of them (almost 100%) with sensitivity 0 (they do not respond to any odor patterns used). This high inactivity seems to be related to the sparse activity shown in Fig. 4.6.

Once known the classification error for different neural pruning strategies, we analyzed the population resulting from the neural pruning for which the minimum classification error was reached. The aim of this was to observe how many neurons

were necessary to reach this minimum and what neuronal sensitivities they had.

In Figure 4.7, we see the initial neural sensitivity of KCs (blue line) and the neural sensitivity for these neurons that achieve the minimum classification error (red line), for the $G \rightarrow S$ pruning method, low noise and the first three overlap degrees showed in Fig. 4.6. For the initial sensitivity, we observe that if overlap degree (std) or connection probability (p_c) increases, the proportion of neurons with high neural sensitivity (generalists) also rises. In the case of the optimal sensitivity, the pruning process leaves just the most specialist neurons (usually only those with neural sensitivity 1). This causes that in the majority of the cases the percentage of inactive neurons in the hidden layer (sensitivity 0) is close to 100%.

4.3 Stimulus space complexity determines the ratio of specialist and generalist neurons

The previous study of the role of specialist and generalist neurons in pattern recognition by neural pruning showed that specialists were more relevant than generalists for odor classification (see Fig. 4.7). It also found that the optimal sensitivity distribution for KCs (the one that achieved the minimum classification error) varied according to the complexity of patterns. These results suggested that the ratio of specialists and generalists in KCs might be an animal adaptation to the complexity of stimuli present in the habitat. To investigate whether the ratio of specialists and generalists changes with the stimuli complexity, we developed a method to create an artificial hidden layer (KCs) with the ratio of these neurons that we want to analyze.

This method, detailed in section 3.2.2.3, calculates the neural sensitivity of the hidden layer (see section 3.2.2.1) and divides these sensitivities equally between specialists (low sensitivities) and generalists (high sensitivities). Therefore, if we have 10 classes, the sensitivities of specialists go from 1 to 5 ($S1-5$) and the ones of generalists from 6 to 10 ($G6-10$). However, sometimes we define specialist and generalist neurons in a more strict way, as specialists can only respond to one pattern class ($S1$) and generalists to all of them ($G10$). Examples of this strict definition of specialists and generalists ($S1/G10$), and the softest definition

mentioned above ($S1-5/G6-10$), as well as intermediate definitions, can be seen in Fig. 4.8 for Gaussian patterns. Once the specialists and generalists of KCs have been separated, we create the new hidden layer using these two sets. This entails that we have to clone some neurons of one type for replacing neurons of the other type. To analyze the odor classification for different ratios of specialists and generalists, we started with a configuration of only generalists and ended with only specialists.

To analyze what is the best ratio of specialists and generalists for pattern recognition according to the complexity level of input patterns, we used different sets of Gaussian patterns (shown in section 3.3.2) in which we have varied the overlap between classes to generate different complexity degrees. The complexity degree was calculated by a measure that links the overlaps between pattern classes (interclass) and inside of them (intraclass) shown in appendix C. Each set of Gaussian patterns have 100 patterns for each of the 10 classes (1000 patterns in total). To compare the results for these patterns and generalize the conclusions, we used the data sets of MNIST digits [115] (10 classes with 100 patterns for each of them) and electronic noses data [184, 225] (6 classes with 100 patterns for each of them).

The following averaged results of this section (see Figs. 4.8-4.12) were obtained from 10 simulations with 5-cross-validation. For these results, we used a connection probability for C matrix, which connects AL to MB, of $p_c = 0.1$, since this value allows a good classification performance, as we saw previously [144, 146]. The connections of W matrix (initialized by random integer values) are subjected to supervised Hebbian learning with negative reinforcement. The probabilities of reinforcement and weakening of weights for this learning are $p_+ = 1$ and $p_- = 0.05$ because of their good performance [90, 147].

4.3.1 Gaussian patterns

As mentioned above, we used Gaussian patterns for analyzing the ratio of specialists and generalists for different complexity levels in the input layer (by increasing the standard deviation of Gaussian distributions and, therefore, the overlap between pattern classes). In Fig. 4.8, we observe that when there is almost no interclass overlap (STD 1 with an overlap of $3.8 \cdot 10^{-6}$, panel (a) in Fig. C.4) and

the complexity is 7.87% the system does not have any problem to classify correctly. The system achieves the classification error 0 for almost any ratio of specialist and generalist neurons and definition of them ($S1/G10$ requires at least 40% of specialist neurons and $S1-2/G9-10$ just 10%).

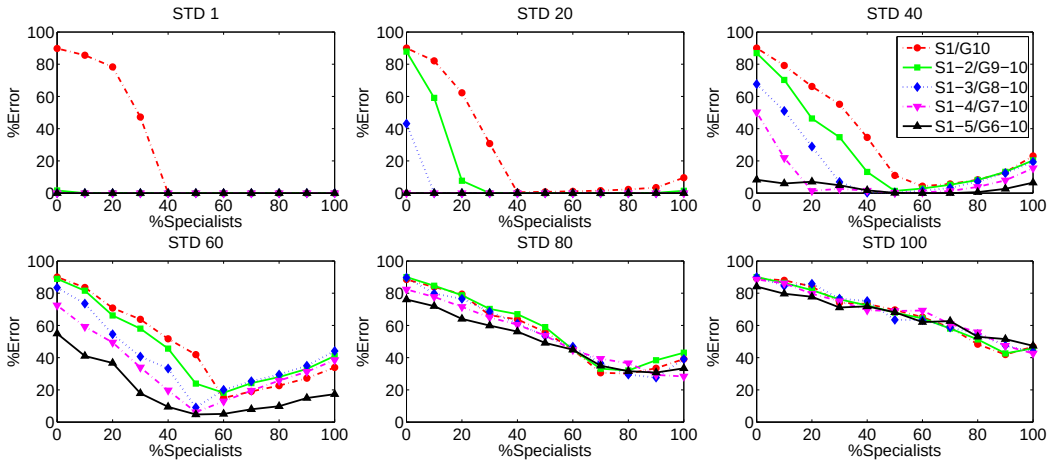


Figure 4.8: Error across the percentage of specialists for Gaussian patterns. We can see that for Gaussian patterns STD 1 and STD 20, the system requires a low neural specialization to minimize the classification error. While for STD 40 and STD 60 both types of neurons, specialists and generalists, are needed (more visible for STD 60). Finally, to STD 80 and STD 100, the system requires mostly specialist neurons to properly classify.

We can see that this behavior starts to change for STD 40, with an interclass overlap of 56.62% and complexity of 28.59%. For this case, the system needs some generalist neurons to achieve the minimum classification error for all boundaries between generalist and specialist neurons. This relevant role of generalist neurons is even clearer when the complexity produced by overlap increases to 62.25%, STD 60. For this intermediate complexity, we can also observe greater variability in the error minimization of the different definitions of generalists and specialists. In this case, it is the definition that covers all neural sensitivities, $S1-5/G6-10$ (see panels (b,c) in Fig. 3.1), the one that reaches the best classification error. However, when the complexity is higher, the relevance of generalist neurons starts to disappear, STD 80 and STD 100 (65.71% and 77.64%), until only the specialist ones are required to achieve the minimum error.

To understand these different behaviors in Fig. 4.8, we suggest focusing on the characteristics of stimuli, Figs. 3.11 and panel (a) in Fig. C.4. For Gaussian patterns with STD 1 we observe that all neurons in the input are specialist and the pattern classes are well-defined, since the interclass overlap is almost 0 and the intraclass one is 92.13%. These patterns are highly separable and highly classifiable and, therefore, they present no difficulty to the system. Although for Gaussian patterns with STD 20 we have a greater generalization of input neurons and a higher interclass overlap (33.84%), the patterns are still easily separable. This situation changes for Gaussian patterns with STD 40. For these patterns, all neurons respond to all classes (generalists) and the interclass overlap is 56.62%. However, there is high variability in the neural activity of these patterns, which provides information about their classes. This entails that even generalist neurons provide useful information for classifying. But when this variability is lower and there is a lot of noise in the patterns, the system no longer differentiates them. This is the case of Gaussian patterns with STD 80 or higher. For these patterns, only specialist neurons are helpful for pattern recognition.

4.3.2 MNIST patterns

For MNIST digits, Fig. 4.9, the system needs at least half of KCs as specialist neurons to achieve the minimum classification error 20.18% for $S1-5/G6-10$ (an error value that is consistent with other studies that use similar computational models [89, 147]). However, if we increase the percentage of specialist neurons more than 50%, we observe that the classification error practically does not change (its maximum value in this region is 20.45%) for all definitions of specialist and generalist (from $S1/G10$ to $S1-5/G6-10$ based on panel (b) in Fig. 3.6). This fact, the invariability of the classification error from 50% of specialists to 100% of them, indicates that the odor classification is accomplished only by specialist neurons, needing at least 50% of these neurons. Thus, we can conclude that for MNIST digits, specialist neurons are the only type of neurons needed to perform a good classification.

This greater relevance of specialist neurons is consistent with what was previously seen for Gaussian patterns. Whenever the complexity is high (85.83% for

MNIST digits, panel (b) in Fig. C.4), only specialist neurons are necessary for classifying. However, there is a substantial difference between the results for MNIST digits and the ones observed in the Gaussian patterns with great complexity. As we can see for MNIST digits, the classification error is 20.18% while for Gaussian patterns with STD 100 is 41.95%. Furthermore, in MNIST digits, Fig. 4.9, there is a stabilization of the classification error, something that does not happen for the Gaussian patterns, Fig. 4.8, since the error decreases as the number of specialists increases. This contrast between patterns with similar complexity levels is due to the difference in the interclass and intraclass overlaps shown in Fig. C.4. In the case that Gaussian patterns had similar values for the overlaps, the relationship between the classification error and the percentage of specialist neurons would be similar to the one observed for MNIST digits, as shown in Fig.14 of [148].

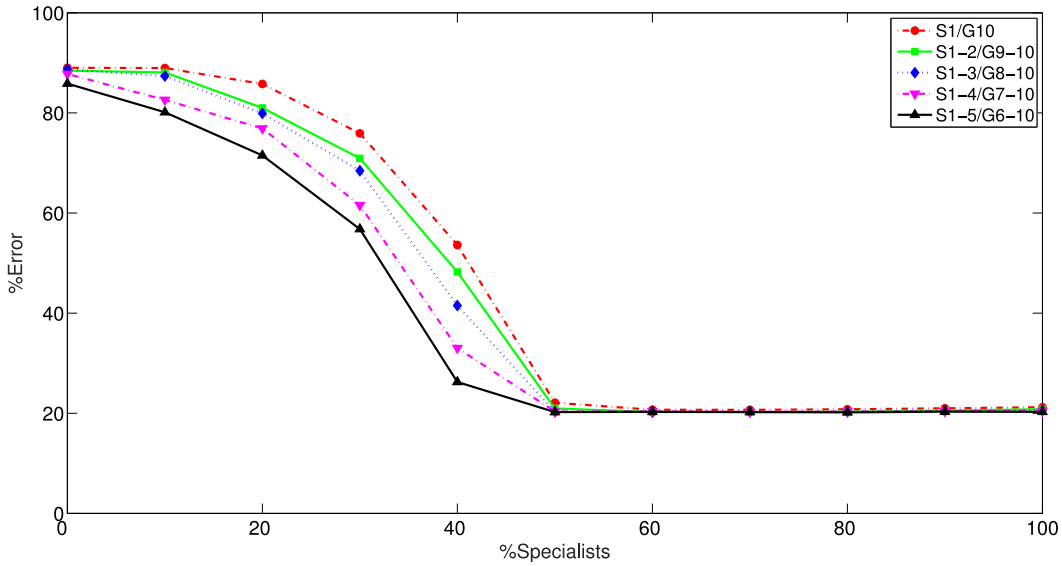


Figure 4.9: Error across the percentage of specialists for MNIST patterns. We can observe that the error minimization is achieved when there is a 50% of specialist neurons and the error varies insignificantly as this percentage increases. So, we can conclude that only specialist neurons are necessary for these patterns.

4.3.3 Electronic noses patterns

For electronic-nose data with a complexity of 49.23% (see panel (c) in Fig. C.4), we observe in Fig. 4.10 that the system achieves the minimum classification error by using specialist and generalist neurons. In fact, the percentage of specialists required to achieve this minimum is between 10% and 20% for all definitions of specialists and generalists. This is consistent with the behavior of Gaussian inputs with similar complexity as the electronic-nose data. Gaussian inputs with STD 60 also require a balance between generalists and specialists to minimize the classification error. In fact, if we look at Fig. 4.8, we see that the classification error obtained by these Gaussian inputs (4.75%) is similar to the one obtained by the electronic-nose data, 5.78% (value obtained by $S1/G6$ and slightly lower than other specialist-generalist definitions). The only different aspect compared to these Gaussian patterns is the error across the percentage of specialists, as the error increases slowly after reaching the minimum (like Gaussian patterns with similar interclass overlap, as STD 40).

4.3.4 Balance of generalist and specialist neurons

After calculating the ratios of generalists and specialists that minimizes the classification error for the different types of patterns, we analyzed the relationship between these ratios and the stimulus complexity. For this purpose, we focused on the percentage of generalists and extracted the minimum and maximum value required for achieving the minimum error. Since we have different definitions of generalist and specialist neurons based on their neural sensitivity (see panels (b,c) in Fig. 3.6), we extracted and averaged these generalist percentages and the minimum error for all of these definitions. Thus, in Fig. 4.11, we can observe the relationship between stimulus complexity and these parameters.

If we compare the panels (a,b) of Fig. 4.11, we can see three different regions. The first one has a great variability between the minimum and maximum percentage of generalist neurons (U). The second one has an intermediate percentage of these neurons (Balance S/G). Finally, for the last one, the percentage of generalist neurons is low (S). The boundaries of these regions are calculated according to the complexities for which the minimum percentage of generalists is 0 (boundaries

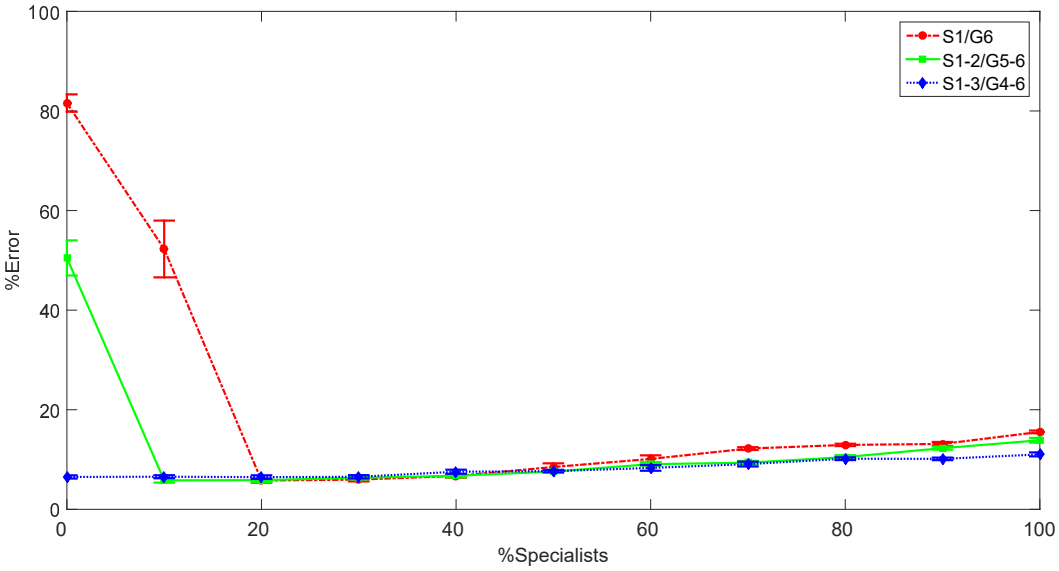


Figure 4.10: Error across the percentage of specialists for electronic-nose inputs. We can observe that the error minimization is achieved for a percentage of specialist neurons between 10% and 20%. This balance between specialists and generalists is consistent with the previous results for Gaussian patterns with similar complexity levels. However, since the classification error grows slowly from 20%, we included error bars in this figure (which represent the standard deviation) in order to show the significance of the results and the relevance of generalists for this dataset.

averaged for different specialist-generalist definitions, panels (b,c) in Fig. 3.1) in order to differentiate for which complexities the generalist neurons are needed for classifying (Balance S/G) and for which ones they are not (U and S). However, if we observe the maximum percentage of generalist neurons, we can see that one of these regions where generalist neurons may not be needed, specialist neurons are also practically irrelevant. This entails that the ratio of these types of neurons needed for a proper classification is unspecific (U) in this region. Therefore, these three regions determine the relevance of generalist and specialist neurons as a function of the complexity of patterns.

If we have a low complexity of patterns (U), the system can reach the classification error 0 (see panel (c) in Fig. 4.11) with a weak specialization of the KC layer (few specialist neurons or soft generalist neurons, see panels (b,c) in Fig. 3.6). When this complexity is intermediate (Balance S/G), the system requires a balance

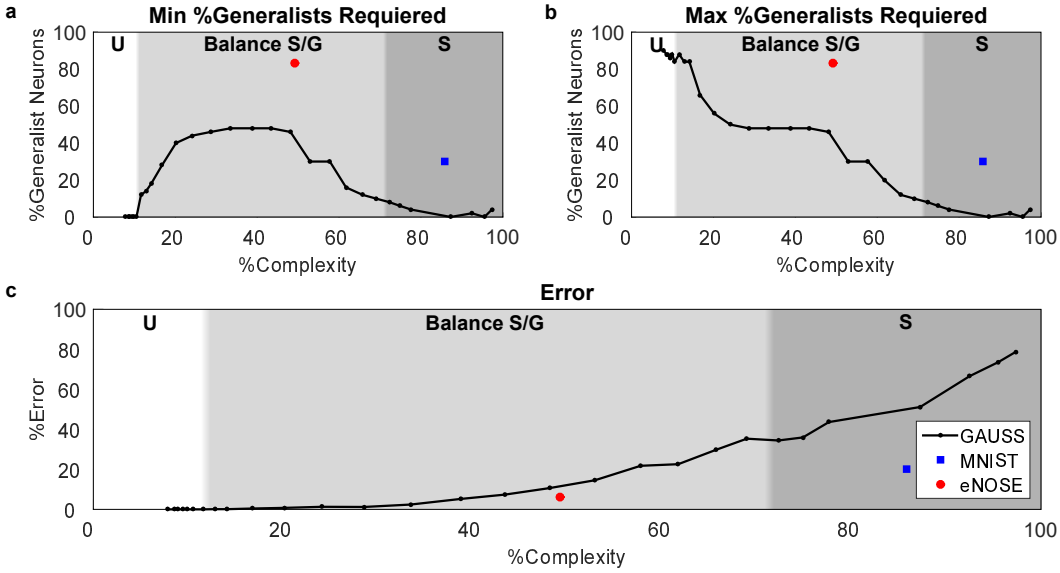


Figure 4.11: Relationship between complexity, the required percentage of generalist neurons and classification error. The top panels show the minimum (panel (a)) and maximum (panel (b)) percentage of generalist neurons needed to achieve the minimum classification error for different Gaussian patterns, MNIST digits and electronic-nose data (eNOSE). The bottom panel (c) shows the relationship between the minimum classification error and the complexity level for these same patterns. Furthermore, we define 3 different regions related to the complexity degree. Firstly, when the complexity is low, there is a great variability between the minimum and maximum percentage of generalist neurons, which entails that the specialist-generalist ratio is unspecific (U). Secondly, when the complexity is intermediate, both types of neurons are needed for classifying (Balance S / G). Finally, when the complexity is high, only specialists are necessary to reach the minimum error (S).

between specialist and generalist neurons (both are relevant). Finally, when the complexity of patterns is high (S), the pattern recognition problem is such that only specialist neurons can minimize the classification error.

Once we have seen the relevance of generalist and specialist neurons for different complexity degrees for Gaussian patterns, it remains to show the percentage of generalist neurons and the classification error that we obtained for each of these regions.

In terms of percentage of generalist neurons (panels (a,b) in Fig. 4.11), we ob-

serve that in the region where its ratio is unspecific (U), the system needs between 0% and 18% of generalist neurons as minimum and between 84% and 90% as maximum. Therefore, we can easily ensure the minimum classification error with almost any percentage of these neurons, such as having half of generalists and half of specialists. For the region where only specialist neurons are relevant (S), the system needs between 0% and 6% of generalist neurons (the minimum and maximum percentages are identical). However, the variability of minimum error is negligible (Fig. 4.9), so we can consider that we do not need any generalist neurons, 0%. In the case of the region where both types of neurons are needed to reach the minimum error (Balance S/G), the system needs between 8% and 48% of generalist neurons as a minimum and between 8% and 66% as a maximum. Although these percentages differ only at the beginning of the region. This implies that the percentage of generalist neurons depends strongly on the complexity of the problem. However, we note that when the complexity is below 50%, the system requires around 50% of generalist neurons. Once we surpass this complexity percentage, the percentage of generalist neurons decreases almost linearly to 8%. This could be the first step to predict the percentage of generalist neurons needed to solve a problem of a certain complexity.

For classification error (panel (c) in Fig. 4.11), we can observe that in the region where the ratio of specialist and generalist neurons is unspecific (U), this error is 0. For the region with the balance between generalist and specialist neurons (Balance S/G), the classification error increases from 0.1% to 34.45%, for Gaussian patterns. This error will continue to rise for the last region (S), where only specialist neurons are needed to achieve the minimum error.

This behavior of Gaussian patterns is also reflected in MNIST digits and electronic-nose data (eNOSE), whose values are displayed with a blue square and red circle respectively in Fig. 4.11. As we can see, MNIST digits are located in the region where only specialist neurons are relevant, while electronic-nose data is placed in the region of balance between generalist and specialist neurons. These locations are consistent with results showed previously for these types of patterns and similar to Gaussian patterns in terms of the relationship between classification error and complexity. A result that also coincides with the ones obtained by LDA and linear SVM classifiers in Fig. C.5. The main difference when compared to Gaussian patterns,

is in terms of the minimum and maximum percentage of generalist neurons. However, if we observe the results for these two patterns, Fig. 4.9 and Fig. 4.10, we see that the minimum classification error varies very slightly for many configurations of generalist and specialist neurons. In the case of MNIST digits, for percentages of generalist neurons between 0% and 50%, the mean error ranges from 20.34% to 20.79%. For electronic-nose data, when the percentage of generalist neurons is from 60% to 80%, the mean error is between 6.04% and 7.92%. Thus, their variations from Gaussian patterns are not so relevant.

4.3.5 The ratio of specialist and generalist neurons in the feature extraction phase determines the odor processing capabilities of the locust olfactory system

As we can infer the ratio of specialist and generalist neurons as a function of the complexity of stimuli, we can do the reverse analysis. We can use neural recordings for estimating the complexity of stimuli that the brain can process. Therefore, we used data from KCs of the locust [122, 175] to calculate the ratio of specialist and generalist neurons, based on the neural responses of 43 neurons for 17 different stimuli [182]. As observed above (see Figs. 4.8-4.10), the system usually achieves the minimum classification error when it has neurons from all possible neural sensitivities ($S1-5/G6-10$ and $S1-3/G4-6$). To divide in half the neural sensitivities shown in the distributions of panel (b) of Fig. 4.12 and Fig.3 of [182], we have taken into account that the maximum sensitivity is in fact 14. Hence, we establish as specialist neurons the ones with neural sensitivity equal or less than 7 and the rest as generalist ones. For the recordings of these 43 KCs and this specialist-generalist definition, we can estimate that the percentage of generalist neurons in the locust is 23.26% (see panel (b) of Fig. 4.12). This ratio involves a complexity of 51.34% according to our calculations (see $GenREC_{KC}$ in panel (c) of Fig. 4.12), which also provides information about the number of differentiable odors by the locust since complexity and capacity are related [181].

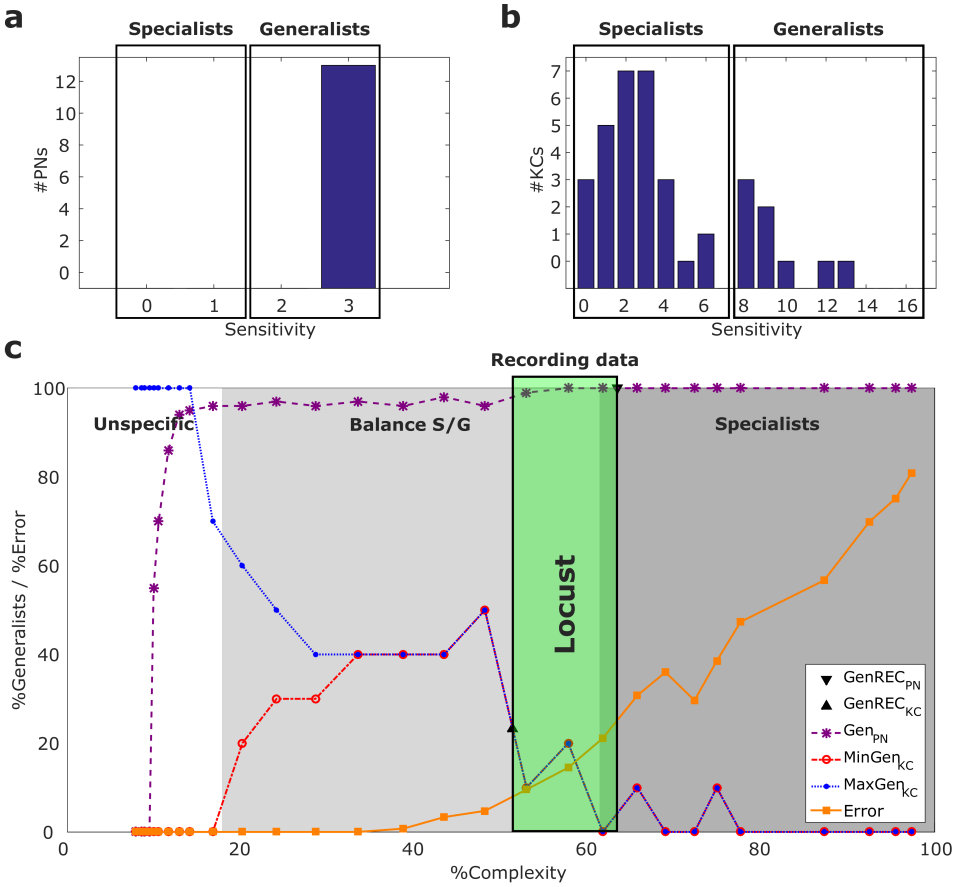


Figure 4.12: Estimation of the accuracy and capacity of the locust olfactory system in odor processing based on the ratios of specialists and generalists observed in neural recordings. We observed previously, that the minimum classification error was achieved by the system when we used neurons from all neural sensitivities and these sensitivities were divided in half between specialists and generalists. Therefore, for PNs recordings [122, 175] (panel (a)) we defined specialists as those neurons with sensitivity from 0 to 1, and generalists as those ones with sensitivities from 2 to 3. In the case of KCs recordings [122, 175] (panel (b)), given that the maximum of different stimuli to which these neurons can respond is 14 [182], we defined specialists as those neurons with neural sensitivity equal or less than 7 and the rest as generalist ones. On the other hand, panel (c) shows the percentage of generalists in the PNs of the locust for our Gaussian patterns and $S1-5/G6-10$ configuration (Gen_{PN}), the minimum and maximum percentages of them estimated in the KCs ($MinGEN_{KC}$ and $MaxGEN_{KC}$) and the minimum classification error for different complexity levels ($Error$). To contrast these results, we provide the percentage of generalists observed in neural recordings for PNs and KCs ($GenREC_{PN}$ and $GenREC_{KC}$). These two values allow us to estimate the stimuli complexity that the locust processes and the classification error of its olfactory system (*Recording data* box).

To contrast these results, we measured the complexity of patterns in PNs, using the recordings of 14 PNs for 3 different odorants [122, 175]. The complexity degree observed for this reduced number of neurons and odors was 63.38% (see $GenREC_{PN}$ in panel (c) of Fig. 4.12) that is not too far from the 51.34% calculated from KCs. This complexity implies that all PNs are generalists, which agrees with the recordings data [175] (see panel (a) of Fig. 4.12) and the value calculated by Gaussian patterns (see GenPN in panel (c) of Fig. 4.12). In addition, from the two complexity values shown, we can estimate that the classification error of the locust olfactory system is between 7.96% and 25.13% (see Recording data box in panel (c) of Fig. 4.12). Thus, we conclude that it is possible to infer the accuracy and capacity of the insect olfactory system in odor processing by analyzing their populations of S/G neurons.

4.3.6 Reproducibility of the balance S/G in the *Drosophila*

We showed above that in a neural network that simulated the locust olfactory system, pattern recognition was influenced by the balance of specialist and generalist neurons. However, these results may not be generalizable to other insects, so we have based on a recent and extensive study on *Drosophila* [37] to test our results using a computational model that simulates its olfactory system. Some of the differences between the locust neural network and the *Drosophila* one are the number of neurons in the antennal lobe (AL) ($\sim 1,000$ [175] vs ~ 250 [37]), the number of Kenyon cells (KCs) in the mushroom body (MB) ($\sim 50,000$ [175] vs $\sim 2,500$ [37]) and the connection probability between the AL and the KCs (~ 0.2 [146, 175] vs ~ 0.01 [37]).

In Fig. 4.13, we can see that when the overlap is less than 28%, the maximum success is achieved for all combinations of specialist and generalist neurons in the KCs. Once this percentage of overlap has been surpassed, the maximum classification success rate is only achieved for a specific balance of these neurons. This balance initially requires a small number of specialists ($10-20\%$), but for overlaps greater than 70% this number increases quickly. This growth causes the neural network of the *Drosophila* finally only needs specialists to achieve the highest clas-

sification success, for input patterns with extremely high overlap ($\sim$ 90%) and, therefore, high complexity.

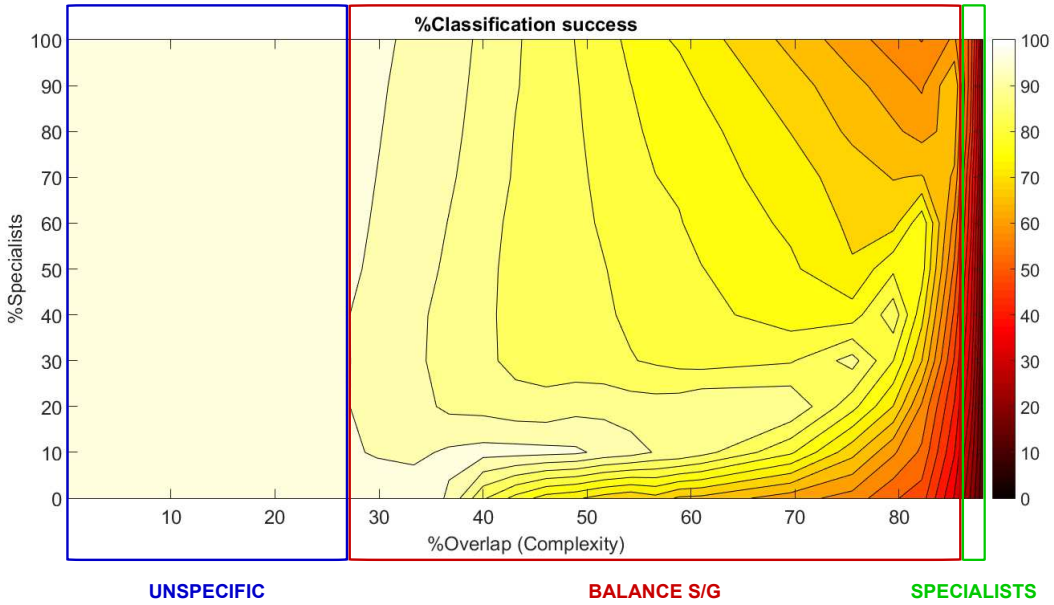


Figure 4.13: Relationship between overlap, the required percentage of specialist neurons and classification success. This picture shows the percentage of specialists required to achieve the maximum classification success as a function of the overlap between patterns. When the overlap is less than 28%, the maximum success is achieved for all combinations of specialists and generalists in KCs. For an overlap from 28% to 90%, the system requires a balance between these two types of neurons to classify correctly. During this period, the number of specialists required by the system increases quickly. Finally, for overlaps higher than 90%, the classification gets worse and the system only needs specialists for improving its performance.

These results are consistent with those observed for locust, Figs. 4.11 and 4.12. The only remarkable difference between both results is that the region of balance between specialists and generalists is greater in *Drosophila*, as well as the percentage of specialist neurons required on it, which is usually lower. This difference may be due to the fact that the connection probability between AL and MB in *Drosophila* ($p_c \sim 0.01$) [37] is much lower than the one estimated for the locust ($p_c \sim 0.2$) [146, 175]. Thus, because the probability that patterns of different classes overlap in KCs is lower for $p_c \sim 0.01$, which entails a decrease in the com-

plexity of the patterns, the number of specialist neurons needed to correctly classify is also lower.

4.4 The neuronal sensitivity of KC does not strongly depend on the number of connections received from PNs

The extensive study on *Drosophila* [37], which we used in section 4.3.6, differentiates KCs according to their incoming connections. Hence, it would be interesting to know whether the randomness of the network that connects PNs to KCs, matrix C (see Fig. 3.1), disappears after the balance between specialists and generalists.

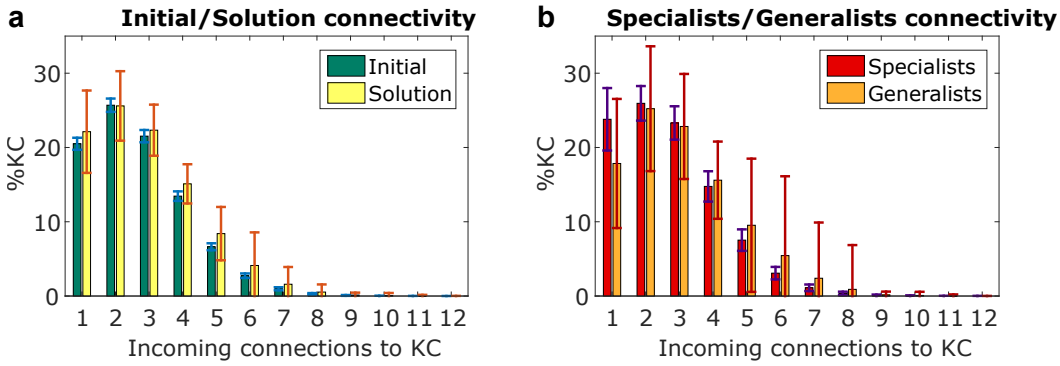


Figure 4.14: Number of connections to Kenyon cells for initial and solution connectivity matrices and specialist and generalist neurons These panels show the mean values for different simulations and overlap degrees, the standard deviations of these values are represented by error bars. Panel (a) shows the connectivity distributions of the initial random matrix and the solution matrix obtained by the optimal balance between specialists and generalists. Panel (b) shows the connectivity distributions of the sets of specialist and generalist neurons extracted from the initial hidden layer (panel (d) of Fig. 3.6).

As shown in panel (a) of Fig. 4.14, the connectivity distributions between the initial matrix C and the solution matrix C' (after the balancing process, see section 3.2.2.3) are similar. The reason for not losing the random structure of connectivity by the balancing process might be the similarity between the connectivity

distributions of the sets of specialists and generalists (panel (b) of Fig. 4.14) extracted from the initial hidden layer (panel (d) of Fig. 3.6). Therefore, when we modify the specialist and generalist populations in the KC layer, subsection 3.2.2, we do not affect the number of connections between PNs and KCs and how they are distributed. Our interpretation is that the neural sensitivity of a neuron is not directly proportional to the number of incoming connections, but to the spatial distribution of stimuli in the input layer of AL.

4.5 Gain control improves the classification performance

For the previous studies on neuronal diversity, we introduced the gain control mechanism observed in AL [167, 168, 235] as simple as possible, by normalizing the input patterns [144, 145, 146, 147, 148, 149]. However, we later decided to study this gain control and perform a bioinspired implementation of it [150]. The aim of developing a computational model that emulates this mechanism is to understand how it works and what advantages for odor classification has the activity regulation of odor patterns in AL.

As we mentioned above in section 3.2.3, the first computational model of this gain control mechanism only took into account the fact that this activity regulation in PNs was produced through the inhibition received by LNs (see panel (a) in Fig. 3.7). Thus, we implemented the lateral inhibition using the following non-linear relation (described previously in Eq. 3.24):

$$PN_j = PN_j - PN_j\beta \left(\sum_{i=1}^{N_{LN}} LN_i - \delta \right), \quad j = 1, \ldots, N_{PN}\,, \tag{4.1}$$

where $\beta \sim 1/\sum_{i=1}^{N_{LN}} LN_i$ is the weight of the inhibitory connection from LNs to PNs, and δ is a threshold. Finally, the non-linear multiplication of $\sum_{i=1}^{N_{LN}} LN_i$ and PN_j assures that PNs with lower activities will remain with lower activities.

Using initially the different concentrations shown in Fig. 3.12 for different sets of Gaussian patterns with 100 and 1000 patterns (10 pattern classes), we observed that this gain control mechanism improved odor classification especially for large

concentrations (Fig. 4.15). For lower concentrations, the difference between using or not gain control is insignificant because the low complexity of patterns for these concentrations makes their classification easy for the system. Finally, if we observe the results obtained for different noise levels, we can conclude that the existence of a gain control mechanism does not prevent the problems caused by noise. This is because when the level of noise increases, the increment of classification error is similar when there is gain control that when there is not.

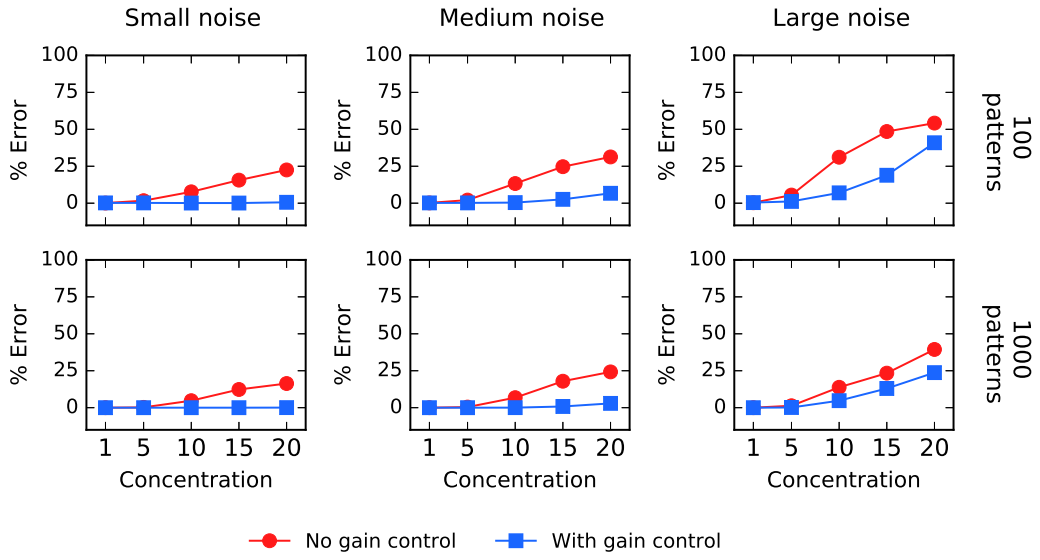


Figure 4.15: Gain control improves classification performance, especially for higher concentrations. All results are the average of 10 simulations for the test data set using supervised learning and 5-fold cross-validation. We show the classification performance when either 100 (top) or 1000 (bottom) different patterns were presented to the neural network. **Left:** With gain control, the classification error rate is insignificant. **Center:** With moderate levels of noise, the gain control mechanism maintains a reasonable performance. **Right:** With a high noise, although the performance drops significantly with the concentration, the network that employed gain control is always 10% more accurate than the one without it.

However, the true potential of the gain control mechanism is to work with different levels of odor concentration at the same time. Hence, in order to quantify this potential, we used datasets of 5000 patterns divided into 10 pattern classes and the 5 concentration levels showed in Fig. 3.12. The performance of a network with

this gain control mechanism (blue line in Fig. 4.16) presented significantly lower classification error rate than a network without gain control (red line in Fig. 4.16), with an improvement of 75%.

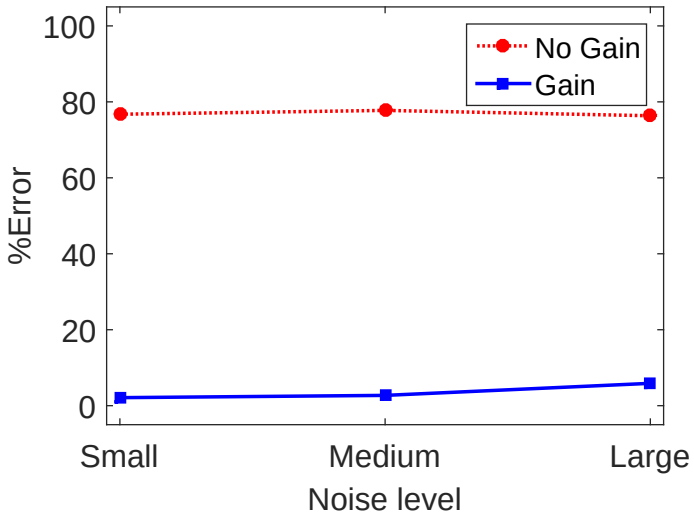


Figure 4.16: Gain control improves the classification performance under wide concentration variability. The classification error using the gain control mechanism is 2% for sets of patterns with small and medium noise and $< 6\%$ for large noise. However, if the gain-control mechanism is not present, the obtained error is around 77%. Thus, the gain control mechanism in AL allows an improvement of 75% in the classification error obtained by MBONs.

However, as we mentioned in section 3.2.3, the previous gain control mechanism described had two major problems from the biological point of view:

- We use the activity of each glomeruli (symbolized by PN_j in equation 4.1 as pre-inhibitory activity) and the strength of the synaptic connection β (which varies depending on the total activity) to emulate the progressive calibration of the activity in PNs through their spatio-temporal dynamics in our spatial model. However, there are not other biological evidence to support them.

- The regulation of activity of PNs did not resemble that observed biologically (panel (d) of Fig.1 in [168]), where the gain control mechanism does not affect very low odor concentration levels.

Hence, to develop a gain control mechanism closer to the biological one, we decided to take into account the existence of two types of LNs: homoLNs and heteroLNs. For this heterogeneous gain control mechanism (see panel (b) in Fig. 3.7), we made the following assumptions: (i) all PNs are uniglomerular (as it happens in the majority of PNs of the *Drosophila*), (ii) homoLNs are connected to all the glomeruli and use the activity received by one of them to inhibit the rest and (iii) heteroLNs are connected to a few glomeruli and use their average activity to inhibit them. This gain control mechanism can be expressed mathematically as follows (described also previously in Eq. 3.25):

$$PN_p^{GL_i} = PN_p^{GL_i} - \alpha Inhib_{homo} - \beta Inhib_{hetero}, \;\; p = 1, \ldots, N_{PN}\,, \tag{4.2}$$

$$Inhib_{homo} = \left(\sum_{j \neq i}^{N_{GL}} \sum_{k=1}^{N_{homoLN^{GL_j}}} homoLN_k^{GL_j} \right)\,, \tag{4.3}$$

$$Inhib_{hetero} = \sum_{l=1}^{N_{heteroLN^{GL_i}}} heteroLN_l^{GL_i}\,, \tag{4.4}$$

where $PN_p^{GL_i}$ is the p-th PN that innervates in the i-th glomerulus, $homoLN_k^{GL_j}$ is the k-th homoLN that receives the odor information of the j-th glomerulus and inhibits the rest and $heteroLN_l^{GL_i}$ is the l-th heteroLN that innervates in the i-th glomerulus (see panel (b) in Fig. 3.7). On the other hand, N_{GL} represents the number of glomeruli, $N_{homoLN^{GL_j}}$ the number of homoLNs that receive the odor information of the j-th glomerulus and $N_{heteroLN^{GL_i}}$ the number of heteroLNs that innervate in the i-th glomerulus. Finally the coefficients $\alpha = 1$ and $\beta = 0.04$ allow regulating the inhibition of the two types of LNs ($Inhib_{homo}$ and $Inhib_{hetero}$) for achieving a suitable gain control.

To compare this gain control mechanism with the previously used ones, we run 10 simulation using datasets of 5000 Gaussian patterns divided into 10 pattern classes and the 5 concentration levels $(5, 25, 50, 75, 100)$ showed in Fig. 3.9. This new gain control mechanism ($Hetero$) is not only more faithful to what was observed through biology (see Fig. 3.8), but also slightly improves (by $\sim 3\%$) the classification error reached for the normalization of the patterns and the use of the previous gain control mechanism ($Homo$) as shown in Fig. 4.17.

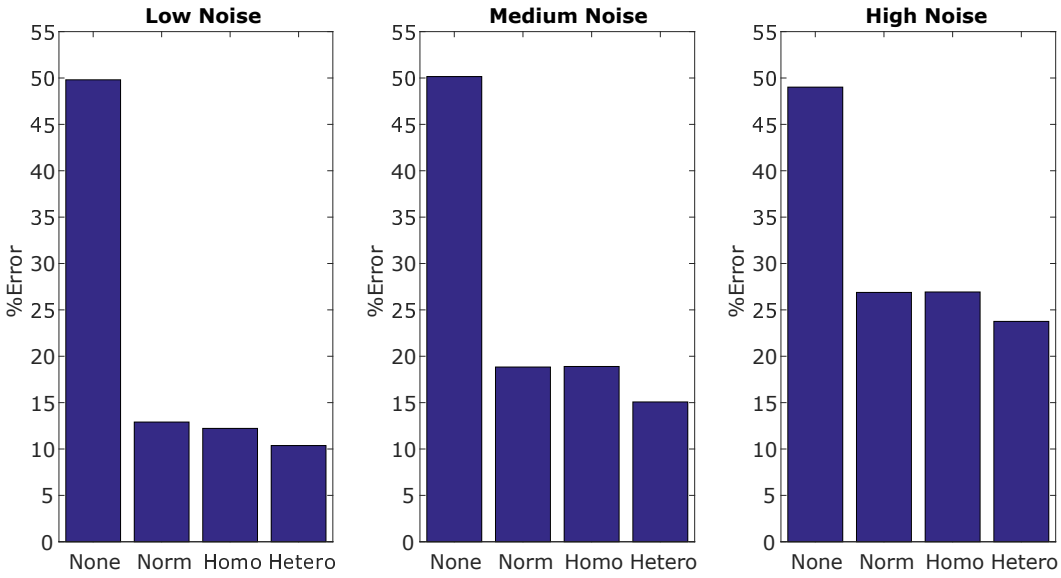


Figure 4.17: Classification error for different gain control mechanisms. As we can observe the heterogeneous gain control mechanism always achieves the minimum classification error. The improvement with respect to the homogeneous gain control mechanisms is $\sim 2\%$ for patterns with low noise, $\sim 4\%$ for medium noise and $\sim 3\%$ for high noise. On the other hand, the homogeneous gain controls obtain similar classification errors that agree with their similar regulation of the input patterns.

CHAPTER

5

Discussions and Conclusions

The purpose of this thesis has been to better understand how the brain processes odor information. The changeful information of odorants (different mixtures, concentrations, environment conditions, etc.) [54] and the high capacity of the olfactory system to process them (for example, a trillion odorants in humans [18] by using 396 different olfactory receptors [158]) make this system interesting from the point of view of machine learning.

To study the olfactory system, we focused on the simple structural organization that this system has in the insects [5, 43, 58, 59, 78, 95, 140, 209], since the main part of this system involved in odor classification can be modeled by a single hidden layer neural network [89]. Using this computational model, we have analyzed a property barely studied and usually not applied in biological models: neural heterogeneity. In particular, we aim to analyze three types of heterogeneity observed in the insect olfactory system in the context of odor discrimination:

- The variability of neural thresholds in KCs [175](see section 4.1).
- The diversity observed in the neural sensitivity of KCs that divides its neural population between specialist and generalist neurons [28, 175, 182, 183, 247] (see sections 4.2-4.4).

- The inhibitory heterogeneity that produces the gain control mechanism observed in AL [65, 193] (see section 4.5).

5.1 Neural threshold heterogeneity improves odor classification

The first study on neural heterogeneity we did was about neural thresholds. Some biological research has proposed that neural thresholds change with time as a function of the input signal that the neuron receives [77, 86]. Furthermore, applied research on artificial noses showed that using heterogeneous detection thresholds for different odorants can improve gas discrimination [35, 47]. Therefore, we have compared the classification error obtained for different data sets, when we use the same threshold for all neurons (homogeneous thresholds) and when we use different thresholds for each of them (heterogeneous). We observed that neural threshold variability can improve classification performance as shown in Figs. 4.1, 4.3 and 4.5. Furthermore, classification errors are lower when the connection probability between AL and MB, p_c is low as well, which is consistent with information maximization criteria provided in [62]. Finally, we proved that success in discrimination is related to the sparse activity documented in the KC layer [175, 214] (see Figs. 4.2-4.4).

5.2 Stimulus space complexity determines the ratio of specialist and generalist neurons

The study on heterogeneous thresholds also showed that these thresholds allowed to differentiate two neuronal populations based on their response to stimuli: specialists and generalists. Specialist neurons are those selective responding to stimuli, while generalists code for multiple stimuli [28]. The role of both kinds of neurons in the olfactory system is still unclear [28, 102]. However, it is suggested that specialist neurons are crucial for discrimination, while generalist neurons play a key role in extracting and discovering common features [236]. Then, we decided to investigate the role of this second neural heterogeneity in pattern recog-

nition problems. For this purpose, we calculated the proper ration of specialists and generalists that minimizes the classification error for different complexities.

We showed (in Fig. 4.11 for the locust and Fig. 4.13 for the *Drosophila*) that if the complexity of patterns is low, the system can reach the minimum classification error with a weak specialization of the system. While, for intermediate complexity, the system will require a balance between specialist and generalist neurons (both are relevant). Finally, when the complexity of patterns is high, the pattern recognition problem is such that only specialist neurons can minimize the classification error. Thus, although specialist neurons are always needed for good classification performance, generalist neurons can be also needed for achieving the global minimum error.

This balance between specialist and generalists neurons does not affect the randomness of the connections between AL and MB in agreement with the biological facts. This fact is due to the similar number of incoming connections of these two types of neurons (Fig. 4.14), which means that the neural sensitivity of KCs seems to be related only to the spatial distribution of stimuli in the AL. Therefore, the regularization of the ratio of specialists and generalists could be applied in randomized neural networks [196] to improve their classification without removing its randomness.

In addition, as we can infer the ratio of specialist and generalist neurons as a function of the complexity of stimuli, we can do the reverse analysis. We can use neural recordings for estimating the complexity of stimuli that the brain can process. Therefore, we used data from KCs of the locust [122, 175] for calculating the ratio of specialist and generalist neurons, based on the neural responses of 43 neurons for 17 different stimuli [182]. We estimated that the percentage of generalist neurons in the locust is 23.26% (see panel (b) of Fig. 4.12). This ratio involves a complexity of 51.34% according to our calculations (see $GenREC_{KC}$ in panel (c) of Fig. 4.12). To contrast these results, we measured the complexity of patterns in PNs, using the recordings of 14 PNs for 3 different odorants [122, 175]. The complexity degree observed for this reduced number of neurons and odors was 63.38% (see $GenREC_{PN}$ in panel (c) of Fig. 4.12) that is not too far from the 51.34% calculated from KCs. Finally, from the two complexity values shown, we can estimate that the classification error of the locust olfactory system is between

7.96% and 25.13% (see Recording data box in panel (c) of Fig. 4.12). Thus, we conclude that it is possible to deduce some capabilities of animal nervous systems by analyzing their populations of S/G neurons.

5.3 The gain control mechanism performed by the heterogeneous inhibition of LNs enhances the odor classification

Finally, we studied the gain control mechanism that keeps the activity of PNs invariant in spite of changes in the odor concentration [167, 168, 235]. This gain control mechanism may be due to the heterogeneity [137] found in the LNs that inhibit the PNs, because two types of neurons are observed: homoLNs, which innervate most if not all glomeruli uniformly, and heteroLNs, which innervate only a few of them [65, 193]. Therefore, simulated this gain control mechanism and analyze its ability to modulate the stimulus signal.

We obtained a gain control mechanism that achieves an input-output relationship similar to the one observed in AL (see Fig. 3.8). Furthermore, we observed that implementing gain control in AL does not only suppress outbursts of activity from input layers but also robustly improves learning in Mushroom Bodies. This improvement is maintained even if we compare the classification errors obtained by our heterogeneous gain control mechanism with those obtained with homogeneous gain control mechanisms, such as pattern normalization (see Fig. 4.17). These results gave a possible computational explanation of why neural heterogeneity can perform a suitable gain control mechanism in the insect olfactory system. Finally, because this gain control can be efficiently implemented, by only adding one extra neural population, this solution may also be of interest to applications. One direct application would be arrays of chemical sensors, the resistance of which increases with the concentration of a given chemical [88, 221].

5.4 Neural heterogeneity is a different approach and a useful tool for solving pattern recognition problems

As we showed during this thesis, our computational model (a randomized neural network [196]) improves odor classification for different kinds of neural heterogeneities. This fact suggests that including these mechanisms in artificial neural networks, chemical sensors and other tools for solving pattern recognition problems can be beneficial.

Sample of the benefits of using these heterogeneities can be seen in some studies. For example, applied research on artificial noses showed that using heterogeneous detection thresholds for different odorants can improve gas discrimination [35, 47]. Furthermore, using threshold variability and combining specialist and generalist systems in artificial neural networks improve their performance [68, 69, 70, 71].

CHAPTER

6

Discusiones y Conclusiones

El propósito de esta tesis ha sido comprender mejor cómo el cerebro procesa la información sobre odorantes procedente del sentido del olfato. La compleja información de los odorantes (diferentes mezclas, concentraciones, condiciones ambientales, etc.) [54] y la gran capacidad del sistema olfativo para procesarlos (por ejemplo, los seres humanos somos capaces de detectar un billón de odorantes [18] utilizando para ello 396 clases de receptores olfativos diferentes [158]) hace que este sistema sea interesante desde el punto de vista del aprendizaje automático.

Para estudiar el sistema olfativo, nos centramos en la simple organización estructural que posee este sistema en los insectos [5, 43, 58, 59, 78, 95, 140, 209], ya que las partes de este sistema involucradas en la clasificación de olores se pueden modelar mediante una red neuronal de una sola capa oculta [89]. Utilizando este modelo computacional, en esta tesis hemos analizado una propiedad generalmente olvidada en los estudios biológicos y en las redes artificiales inspiradas en estos: la heterogeneidad neuronal. En particular, nuestro objetivo es analizar tres tipos de heterogeneidad observada en el sistema olfativo de los insectos en el contexto de la discriminación de olores:

- La variabilidad de los umbrales neurales en las KCs [175] (ver sección 4.1).

- La diversidad observada en la sensibilidad neuronal de las KCs, que permite dividir esta población neuronal en neuronas especialistas y generalistas [28, 175, 182, 183, 247] (vea las secciones 4.2-4.4).
- La heterogeneidad inhibitoria responsable del mecanismo de control de ganancia observado en el AL [65] (ver sección 4.5).

6.1 La heterogeneidad del umbral neuronal mejora la clasificación de odorantes

El primer estudio sobre heterogeneidad neuronal que llevamos a cabo fue sobre los umbrales neuronales. Algunos estudios biológicos sugieren que los umbrales neuronales cambian con el tiempo en función de la señal de entrada que recibe la neurona [77, 86]. Además, una investigación aplicada sobre narices artificiales mostró que el uso de umbrales de detección heterogéneos para diferentes odorantes puede mejorar la discriminación de los gases [35, 47]. Por lo tanto, hemos comparado el error de clasificación obtenido para diferentes conjuntos de datos, cuando utilizamos un mismo umbral para todas las neuronas (umbrales homogéneos) y cuando utilizamos diferentes umbrales para cada una de ellas (heterogéneos). Observamos que la variabilidad del umbral neuronal puede mejorar la capacidad de clasificación de odorantes del sistema, tal como se muestra en las Figs. 4.1, 4.3 y 4.5. Además, los errores de clasificación son menores cuando la probabilidad de conexión entre AL y MB (p_c) también es baja, lo que es consistente con los criterios de maximización de la información mostrados en [62]. Finalmente, probamos que el éxito en la discriminación está relacionado con la escasa actividad observada en las KCs [175, 214] (Figs. 4.2-4.4).

6.2 La complejidad del estímulo determina la proporción de neuronas especialistas y generalistas

El estudio sobre umbrales heterogéneos también nos mostró que, a través de estos, se podían diferenciar dos poblaciones neuronales en función de su respuesta a los estímulos: especialistas y generalistas. Las neuronas especialistas son aquellas

que responden de manera selectiva a los estímulos, mientras que las generalistas se activan para múltiples estímulos [28]. El papel de ambos tipos de neuronas en el sistema olfativo aún no está claro [28, 102]. Sin embargo, se ha sugerido que las neuronas especialistas son cruciales para la discriminación, mientras que las neuronas generalistas desempeñan un papel clave en la extracción de características comunes [236]. Por tanto, decidimos investigar el papel de esta segunda heterogeneidad neuronal para el reconocimiento de patrones. En concreto, estimamos cuál es la proporción adecuada de este tipo de neuronas para la minimización del error de clasificación para diferentes niveles de complejidad en los patrones de entrada.

Los resultados para la langosta de tierra (ver Fig. 4.11) y la mosca de la fruta (ver Fig. 4.13) muestran que cuando la complejidad de los patrones es baja, el sistema puede alcanzar el error mínimo de clasificación para una escasa especialización del sistema, mientras que, para una complejidad intermedia, el sistema requerirá un balance concreto entre neuronas especialistas y generalistas (ambas son relevantes). Finalmente, cuando la complejidad de los patrones es alta, el problema del reconocimiento de patrones es de tal dificultad que sólo las neuronas especialistas pueden minimizar el error de clasificación. Por lo tanto, aunque las neuronas especialistas siempre son necesarias para una buena clasificación, las neuronas generalistas también pueden ser necesarias para lograr el error mínimo global.

Este balance entre las neuronas especialistas y generalistas no afecta a la aleatoriedad de las conexiones entre AL y MB, conservando así las propiedades de la red original. Este hecho se debe a que la cantidad de conexiones recibidas por estas dos clases de neuronas es similar (Fig. 4.14), lo que significa que la sensibilidad neuronal de las KCs parece estar relacionada únicamente con la distribución espacial de los estímulos en el AL. Por lo tanto, la regularización del ratio de especialistas y generalistas podría aplicarse en redes neuronales aleatorias [196] para mejorar su clasificación sin eliminar su aleatoriedad.

Además, de igual manera que podemos estimar la proporción de neuronas especialistas y generalistas en función de la complejidad de los estímulos, podemos hacer el análisis inverso: podemos usar grabaciones neuronales para estimar la complejidad de los estímulos que el cerebro puede procesar. En consonancia con

esta idea, utilizando datos de KCs de la langosta de tierra [122, 175] podemos calcular la proporción de neuronas especialistas y generalistas en el sistema biológico, basándonos en las respuestas neuronales de 43 neuronas a 17 estímulos diferentes [182]. Mediante estos datos, hemos estimados que el porcentaje de neuronas generalistas en la langosta terrestre es del 23.26% (ver el panel (b) de la Fig. 4.12). Este porcentaje de generalistas implica una complejidad del 51.34% según nuestros cálculos (ver $GenREC_{KC}$ en el panel (c) de la Fig. 4.12). Para contrastar estos resultados, medimos la complejidad de los patrones en las PNs, utilizando las grabaciones de 14 PNs para 3 odorantes diferentes [122, 175]. El grado de complejidad observado para este número reducido de neuronas y olores fue del 63.38% (ver $GenREC_{PN}$ en el panel (c) de la Fig. 4.12) el cual no dista tanto del 51.34% calculado a partir de las KCs. Finalmente, a partir de estos dos valores de complejidad, podemos estimar que el error de clasificación del sistema olfativo de la langosta terrestre está entre el 7.96% y el 25.13% (ver "Recording data box" en el panel (c) de la Fig. 4.12). Por lo tanto, concluimos que es posible deducir algunas capacidades de los sistemas biológicos de procesamiento de información analizando sus poblaciones de neuronas especialistas y generalistas.

6.3 El mecanismo de control de ganancia realizado por la inhibición heterogénea de las LNs mejora la clasificación de olores

Finalmente, queríamos estudiar el mecanismo de control de ganancia que mantiene invariante la actividad de las PNs a pesar de los cambios en la concentración del olor [168]. Este mecanismo de control de ganancia puede deberse a la heterogeneidad [137] encontrada en las LNs que inhiben las PNs, ya que se observan dos tipos diferentes de neuronas: las homoLNs, que inervan a la mayoría de los glomérulos de manera uniforme, y las heteroLNs, que inervan solamente a algunos de ellos [65]. Por lo tanto, hemos intentado simular este mecanismo de control de ganancia y analizar su capacidad para modular la señal del estímulo.

El mecanismo de control de ganancia finalmente obtenido permite emular la relación de actividad de entrada-salida observada en las PNs (ver Fig. 3.8). Además,

observamos que la implementación de este control de ganancia en el AL también mejora la clasificación de los olores en el MB, obteniendo mejores resultados para este mecanismo de control de ganancia heterogéneo incluso cuando lo comparamos con mecanismos homogéneos, tales como la normalización de los patrones (ver Fig. 4.17). Finalmente, debido a que este control de ganancia puede implementarse fácilmente, sólo agregando una población neuronal adicional, esta solución también puede ser de interés para ciertas aplicaciones como su uso en sensores químicos, cuya resistencia aumenta con la concentración de un producto químico determinado [88, 221].

6.4 La heterogeneidad neuronal es un enfoque diferente y una herramienta útil para resolver problemas de reconocimiento de patrones

Como mostramos durante esta tesis, nuestro modelo computacional (una red neuronal aleatoria [196]) mejora la clasificación de los odorantes para las diferentes clases de heterogeneidad neuronal estudiadas. Este hecho nos lleva a sugerir que incluir estos mecanismos en redes neuronales artificiales, sensores químicos y otras herramientas para resolver problemas de reconocimiento de patrones puede ser beneficioso.

En algunos estudios se puede ver una muestra de los beneficios de usar estas heterogeneidades. Por ejemplo, la investigación aplicada en narices artificiales determinó que el uso de umbrales de detección heterogéneos para diferentes odorantes puede mejorar la discriminación de gases [35, 47]. Además, el uso de la variabilidad del umbral y la combinación de sistemas especializados y generalistas en redes neuronales artificiales mejora el funcionamiento de estos sistemas [68, 69, 70, 71].

CHAPTER

7

Future work

The knowledge and results obtained during this thesis mark the starting point for many other researches, which are currently under development.

In the near future, we have three investigations in progress, which are listed below according to how close we are from publishing new results:

1. As we discussed in section 3.2.1.2, we are still working on an algorithm for neural threshold selection by gradient descent. This algorithm detailed here in B.2.2, has been updated and improved, being the most current version published in [124].

2. The results of the study of the heterogeneous inhibition of the LNs in the gain control mechanism in AL (shown in section 4.5) have been the last ones presented in this thesis and as preliminary data. However, it is necessary to analyze the heterogeneous gain control model shown (panel (b) in Fig. 3.7) for different data sets (besides Gaussian patterns) and to compare it with other types of inhibitory structures.

3. Finally, we are working on a complexity measure based on the relationship of the inter and intraclass overlaps that improves the one shown in appendix C for the study of specialist and generalist neuron populations (see

section 3.2.2). This new complexity measure seeks to incorporate factors linked to the difficulty level of pattern recognition problems (such as the number of classes and the number of patterns per class), while maintaining the low computational cost of the measure shown here (in equation C.6).

CHAPTER

8

Publications

JOURNAL PUBLICATIONS

A. Montero, R. Huerta, F.B. Rodriguez.
"Stimulus space complexity determines the ratio of specialist and generalist neurons during pattern recognition".
Journal of the Franklin Institute, 355(5), 2951-2977 (2018). Elsevier.
Q1. ***JIF:3.576***.

A. Montero, R. Huerta, F.B. Rodriguez.
"Regulation of specialists and generalists by neural variability improves pattern recognition performance".
Neurocomputing, 151, 69-77 (2015). Elsevier.
Q1. ***JIF:2.392***.

CONFERENCE PUBLICATIONS

A. Montero, J. Lopez-Hazas, F.B. Rodriguez.
"Input Pattern Complexity Determines Specialist and Generalist Populations in Drosophila Neural Network".
International Conference on Artificial Neural Networks, (2018) 296-303. Springer, Cham.

J. Lopez-Hazas, A. Montero, F.B. Rodriguez.
"Strategies to Enhance Pattern Recognition in Neural Networks Based on the Insect Olfactory System".

International Conference on Artificial Neural Networks, (2018) 468-475. Springer, Cham.

A. Montero, T. Mosqueiro, R. Huerta, F.B. Rodriguez.
"Exploring a Mathematical Model of Gain Control via Lateral Inhibition in the Antennal Lobe".
International Work-Conference on Artificial Neural Networks, 10305, 317-326 (2017). Lecture Notes in Computer Science, Springer.

A. Montero, R. Huerta, F.B. Rodriguez.
"Specialist Neurons in Feature Extraction Are Responsible for Pattern Recognition Process in Insect Olfaction".
Artificial Computation in Biology and Medicine, 58-67 (2015). Lecture Notes in Computer Science, Springer.

A. Montero, R. Huerta, F.B. Rodriguez.
"Neural Trade-Offs among Specialist and Generalist Neurons in Pattern Recognition".
Engineering Applications of Neural Networks, 71-80 (2014). Communications in Computer and Information Science, Springer.

A. Montero, R. Huerta, F.B. Rodríguez.
"Neuron Threshold Variability in an Olfactory Model Improves Odorant Discrimination".
Natural and Artificial Models in Computation and Biology, 16-25 (2013). Lecture Notes in Computer Science, Springer.
Award: Best paper in Biocomputation. *Amount: 500€.*

CHAPTER

9

Conferences

2018 Oral presentation *"Exploring a Mathematical Model of Gain Control via Lateral Inhibition in the Antennal Lobe"* at the 27th International Conference on Artificial Neural Networks (ICANN).
CORE B

Oral presentation *"Strategies to Enhance Pattern Recognition in Neural Networks Based on the Insect Olfactory System"* at the 27th International Conference on Artificial Neural Networks (ICANN).
CORE B

Poster presentation *"The ratio of specialist and generalist neurons in the feature extraction phase determines the odor processing capabilities of the locust olfactory system"* at the 27th Computational Neuroscience Meeting (CNS).

Poster presentation *"Regulation of neural threshold in Kenyon cells through their sparse condition improves pattern recognition performance"* at the 27th Computational Neuroscience Meeting (CNS).

9. CONFERENCES

2017 Poster presentation *"Gain control mechanism based on lateral inhibition of antennal lobe improves pattern recognition performance under wide concentration variability"* at the 26th Computational Neuroscience Meeting (CNS).

Oral presentation *"Exploring a Mathematical Model of Gain Control via Lateral Inhibition in the Antennal Lobe"* at the 14th International Work-Conference on Artificial Neural Networks (IWANN).
CORE B

2015 Oral presentation *"Specialist Neurons in Feature Extraction Are Responsible for Pattern Recognition Process in Insect Olfaction"* at the 6th International Work-conference on the Interplay between Natural and Artificial Computation (IWINAC).

2014 Oral presentation *"Neural Trade-Offs among Specialist and Generalist Neurons in Pattern Recognition"* at the 15th Engineering Applications of Neural Networks (EANN).
CORE C.

2013 Oral presentation *"Neuron Threshold Variability in an Olfactory Model Improves Odorant Discrimination"* at the 5th International Work-conference on the Interplay between Natural and Artificial Computation (IWINAC).

CHAPTER

10

Research Internships

2014 BioCircuits Institute.
University of California, San Diego (UCSD).
Advisor: Ramón Huerta
Duration: 89 days.
Internship aim: Study the role of generalist and specialist neurons, observed in the olfactory system of insects, in the odor classification.

2013 BioCircuits Institute.
University of California, San Diego (UCSD).
Advisor: Ramón Huerta
Duration: 82 days.
Internship aim: Mathematical development of a learning algorithm inspired by the olfactory system of insects, using the neural heterogeneity and sparseness condition observed in it.

APPENDIX

A

Glossary of acronyms

AC: Adenylate Cyclase
AL: Antennal Lobe
APL: Anterior Paired Lateral Neuron
ATP: Adenosine Triphosphate
cAMP: Cyclic Adenosine Monophosphate
GDP:Guanosine Diphosphate
GGN: Giant GABAergic Neuron
GPCR: G-Protein-Coupled Receptor
GTP: Guanosine Triphosphate
iACT: Inner Antenno-Cerebral Tract
KC: Kenyon Cell
heteroLN: Heterogeneous Local Neuron
homoLN: Homogeneous Local Neuron
lACT: Lateral Antenno-Cerebral Tract
LH: Lateral Horn
LN: Local Neuron
mACT: Middle Antenno-Cerebral Tract
MB: Mushroom Body
MBON: Mushroom Body Output Neuron

A. APPENDIX: GLOSSARY OF ACRONYMS

OBP: Odorant-Binding Protein
OR: Olfactory Receptor
ORN: Olfactory Receptor Neuron
PN: Projection Neuron
SLN: Single-Hidden-Layer Neural Network

APPENDIX

B

Threshold Selection

To study the variability observed in neural thresholds [2, 175], we compared the classification error obtained when we used different thresholds for each neuron of the KC layer (heterogeneous thresholds) with the one obtained when we used a unique threshold for all of them (homogeneous ones). The selection of threshold values to perform this comparison was made through two methods: exhaustive search and gradient descent algorithm.

In section 1, we described the algorithms used for the exhaustive search of threshold values. This search used the total stimulation received by a neuron for a given odorant, which we called limit threshold (see section 3.2.1.1) and it is the minimum threshold value that prevents the neuron from responding to that odorant. We developed a unique algorithm for the calculation of the homogeneous thresholds and three algorithms for the heterogeneous ones. While the first two algorithms of heterogeneous thresholds aim to compare their results with homogeneous ones (see section 4.1.1), the last algorithm was used to study the specialist and generalist neurons of the KCs (see sections 4.2- 4.4), once we concluded that the heterogeneity in neural thresholds improves the odor classification.

Section 2 details the mathematical reasoning for the gradient descent algorithm. This algorithm had two different approaches. Initially, the equation used had four terms: the first to calculate the optimal neural thresholds of KCs, the second to

minimize the classification error, the third to reduce the neural activity of KCs and the fourth to regulate the weights of the matrix W. Although with this equation we obtained good results for simple problems (shown in section 4.1.2), it generally presented several problems such as a high computational cost and difficulties to escape from local minima. Thus, we decided to simplify it in a two-term equation: one to minimize the classification error and the other to reduce the neural activity of KCs. We are currently refining this learning, whose preliminary results can be found in [123, 124].

B.1 Threshold selection method by exhaustive search

As we mentioned above, the calculation of homogeneous and heterogeneous thresholds using exhaustive search is based on the concept of limit threshold (the minimum threshold value which prevents that the neuron fires for a certain stimulus). The algorithms to calculate these types of thresholds are detailed below.

B.1.1 Homogeneous thresholds

To calculate the homogeneous thresholds of each layer, we set the network weights, C and W, and obtain the limit threshold matrix of the hidden layer. We take the minimum and maximum of this matrix and use these values, and those among them, as thresholds for the hidden layer (Algorithm.1, lines 1-7). The goal is to obtain the minimum classification error for each threshold and the spike rate for the minimum value.

In order to achieve this minimum classification error, we obtain the limit threshold matrix for the output layer, of each hidden layer threshold. Using the minimum, maximum and the values among them of this matrix, we modify the weights of W by Hebbian learning and calculate the classification error for all possible combination of threshold values. We store in an array the minimum classification error for a hidden layer threshold and the spike rate linked to it (Algorithm.1, lines 8-24).

Algorithm 1 Homogeneous thresholds

Set C and W as a function of p_c and p_w
$\Theta = C^t X$ //Limit threshold matrix for the hidden layer
$\theta_{min} = min(\Theta)$ //minimum matrix Θ
$\theta_{max} = max(\Theta)$ //maximum matrix Θ
$N = \theta_{max} - \theta_{min} + 1$ //number of thresholds
$error[N] = 1$ //vector that stores the minimum error for each θ
$spikes[N] = 0$ //vector that stores the spike rate linked to the minimum error of each θ
for $n = 0 \rightarrow N - 1$ **do**
    $\theta = \theta_{min} + n$
    $Y = C^t X - \theta > 0$
    $E = W^t Y$ //Limit threshold matrix for the output layer
    $\varepsilon_{min} = min(E)$ //minimum matrix E
    $\varepsilon_{max} = max(E)$ //maximum matrix E
    $M = \varepsilon_{max} - \varepsilon_{min} + 1$
    for $m = 0 \rightarrow M - 1$ **do**
        $\varepsilon = \varepsilon_{min} + m$
        $Z = W^t Y - \varepsilon > 0$
        $HebbianLearning(z, y, w)$
        if $error < error(\theta)$ **then**
            $error(\theta) = error$
            $spikes(\theta) = spikes$
        end if
    end for
end for

B.1.2 Heterogeneous thresholds

We have used three different methods for heterogeneous threshold selection. First, we calculated the threshold values based on the percentage of odorants for which the neurons would not fire [144]. Second, since the previous method only worked correctly for the primary patterns 3.3.1, we decided to set the threshold values according to the type of neuron based on the variance of its limit threshold

distribution (specialist or generalist) [146]. Finally, for the rest of the thesis we used the heterogeneous thresholds to increase the neural sensitivity of the KCs and, thus, be able to better study the specialist and generalist neurons.

B.1.2.1 Heterogeneous thresholds by excluding a percentage of odorants

Algorithm 2 Heterogeneous threshold selection by excluding a percentage of odorants

1: Set C and W as a function of p_c and p_w
2: $\Theta = C^t X$ //Limit threshold matrix for the hidden layer
3: $N = 100$
4: $error[N+1] = 1$
5: **for** $n = 0 \rightarrow N$ **do**
6: $\theta = percentile(\Theta, n)$ //Value that makes the neurons do not fire for a n percentage of odorants
7: $Y = C^t X - \theta > 0$
8: $E = W^t Y$ //Limit threshold matrix for the output layer
9: **for** $m = 0 \rightarrow N$ **do**
10: $\varepsilon = percentile(E, n)$ //Value that makes the neurons do not fire for a n percentage of odorants
11: $Z = W^t Y - \varepsilon > 0$
12: $HebbianLearning(Y, Z, W)$
13: **if** $error < error(n+1)$ **then**
14: $error(n+1) = error$
15: **end if**
16: **end for**
17: **end for**

We set the network weights, C and W, and obtain the limit threshold matrix of the hidden layer. Using these limit thresholds we make the neurons fire at a certain percentage of odorants that go from 0 to 100 (Algorithm.2, lines 1-6). Then we perform the same process to calculate the thresholds of the output layer (Algorithm.2, lines 7-10).

For each combination of threshold values, we start the Hebbian learning, which will update the weights of W. Once the weights have been learned, we calculate the classification error. The minimum classification error obtained for all possible thresholds will be selected as the final error of the system (Algorithm.2, lines 11-17).

B.1.2.2 **Heterogeneous thresholds based on the variance of the limit threshold distribution**

We set the C and W weights and calculate the limit threshold matrix for the hidden layer. Using this matrix, we obtain the distribution of limit thresholds for each hidden layer neuron (Algorithm.3, lines 1-4). Using the variance of these distributions of limit thresholds, we labeled the neurons as generalists or specialists [182, 236]. Generalist neurons should have a low variance of limit thresholds, since these neurons are similarly stimulated by all odorants, while specialist neurons should have a greater variance because they are more sensitive to a certain odorant (Fig. 3.3). This reasoning is based on the fact that a greater variability of limit thresholds allows the neuron to select a threshold value that makes it selective to certain odorants.

Because there is not a certain variance that allows labeling the neurons as specialists and generalists, we calculate the variance of limit thresholds for each neuron and sort these values from lowest to highest. Subsequently, starting with the lowest variance values, we are going to label a percentage of neurons as generalists from 0 to 100. This process allows us to consider different ratios of generalists and specialists (Algorithm.3, lines 5-8).

Once the neurons are labeled, the neural thresholds are calculated (Algorithm.3, line 9). Because it was suggested that only specialist neurons are necessary for odor discrimination, we assigned the maximum limit threshold to generalist neurons with the objective of "pruning" them. On the other hand, in the case of specialist neurons, we assumed that, due to the great variability of these neurons, their distribution could be bimodal. This entails that there is a minimum value between the two different modes. Hence, we search this minimum value, making the neuron responds only to a few odorants (specialist) since the second mode is usually inferior to the first one. The threshold value linked to this minimum value establishes the

B. APPENDIX: THRESHOLD SELECTION

Algorithm 3 Heterogeneous thresholds

1: Set C and W as a function of p_c and p_w
2: $\Theta = C^t X$ //Limit threshold matrix for the hidden layer
3: $N = 100$
4: $error[N+1] = 1$
5: $var = variance(\Theta)$ //threshold variance for each neuron
6: **for** $n = 0 \rightarrow N$ **do**
7: $p = percentile(var, n)$ //variance value for which a n percentage of neurons is below
8: $label\_neuron(p)$ //it labels using p each neuron as generalist or specialist
9: $\theta = threshold\_selection$ //each threshold is assigned depending on the maximum limit threshold for generalist neurons or α_θ for specialist neurons
10: $E = W^t Y$ //Limit threshold matrix for the output layer
11: $var = variance(E)$
12: **for** $m = 0 \rightarrow N$ **do**
13: $p = percentile(var, m)$ //variance value for which a m percentage of neurons is below
14: $label\_neuron(p)$ //it labels using p each neuron as generalist or specialist
15: $\varepsilon = threshold\_selection$ //each threshold is assigned depending on the maximum limit threshold for generalist neurons or α_ε for specialist neurons
16: $HebbianLearning(z, y, w)$
17: **if** $error < error(n+1)$ **then**
18: $error(n+1) = error$
19: **end if**
20: **end for**
21: **end for**

specialization coefficient α of the neuron, the percentage of odorants to which a neuron stops responding to. This method has its limitations, but in many cases it allows us to achieve the neuron's appropriate specialization (Fig. 3.3).

Finally, we perform the same process to calculate the thresholds of the output layer (Algorithm.2, lines 7-10). For each combination of threshold values, we start

the Hebbian learning, which will update the weights of W. Once these weights have been learned, we calculate the classification error. The minimum classification error obtained for all possible thresholds will be selected as final error of the system (Algorithm.3, lines 10-20).

B.1.2.3 **Heterogeneous thresholds for increasing neural sensitivity**

We set the network weights, C and W, and obtain the limit threshold matrix of the hidden layer. Using these limit thresholds, we increase the neural sensitivity of the hidden layer by making the neurons fire randomly to a certain percentage of odorants that go from 0 to 100 (Algorithm.4, lines 1-10).

Algorithm 4 Heterogeneous threshold selection by excluding a percentage of odorants

1: Set C and W as a function of p_c and p_w
2: $\Theta = C^t X$ //Limit threshold matrix for the hidden layer
3: $N = 100$
4: $N_{KC} = 50000$
5: $M = N_{KC}/N$ //Subdivisions of KCs for each percentage of odorants for which neurons would not fire
6: $id = randperm(N_{KC})$ //Vector for the random selection of KCs
7: $error = 1$
8: **for** $n = 1 \rightarrow N$ **do**
9: $\quad \theta(id((n-1) * M + 1 : n * M)) = percentile(\Theta, n-1)$ //Value that makes the neurons do not fire for a $n-1$ percentage of odorants
10: **end for**
11: $Y = C^t X - \theta > 0$
12: $Z = W^t Y - \varepsilon > 0$
13: $HebbianLearning(Y, Z, W)$
14: **if** $error < error(n+1)$ **then**
15: $\quad error(n+1) = error$
16: **end if**

After this process, we calculate the values of the hidden and output layer and start the Hebbian learning, which will update the weights of W. Once these weights

have been learned, we calculate the classification error (Algorithm.4, lines 11-16).

B.2 Threshold selection method by gradient descent

To study the variability of neural thresholds, we also used a gradient descent algorithm to calculate the homogeneous and heterogeneous thresholds and the weights of matrix W (which represent the connections between KCs and MBONs where the odor learning is produced). This algorithm had two different approaches.

B.2.1 First approach

Initially, we proposed an equation with four terms: the first one to calculate the optimal neural thresholds of KCs based on the ELM algorithm [87], the second one to reduce the classification error, the third one to reduce the neural activity of KCs (given the sparseness condition of these in nature [175]) and the fourth one to regulate the weights of matrix W:

$$E = \alpha \left|Y^* - Y\right|^2 + \beta \left|T - Z\right|^2 + \lambda \left|Y^*\right| + \mu \frac{|W|^2}{2} \tag{B.1}$$

where Y^* is the target hidden layer matrix, Y is the hidden layer matrix, T is the target output matrix, Z is the output matrix and W is the connectivity matrix between Y and Z. Every term is weighted according to their coefficients, α for threshold selection, β for classification error, λ for the sparse neural activity and μ for weight regulation.

The activation function for the neurons of matrices Y and Z is sigmoidal, while for the matrix Y^* it is obtained by means of the ELM algorithm:

$$y_{jp} = \frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)}\right)} \tag{B.2}$$

$$z_{kp} = \frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)} \tag{B.3}$$

$$y^*_{jp} = \sum_{k=1}^{N_{OUT}} w^{\dagger}_{kj} \left(\frac{-log\left(\frac{1}{t_{kp}} - 1\right)}{\gamma} + \varepsilon_k\right) \tag{B.4}$$

Being able to express this last equation in the following way:

$$\tau_{kp} = \frac{-log\left(\frac{1}{t_{kp}} - 1\right)}{\gamma} \tag{B.5}$$

$$y^*_{jp} = \sum_{k=1}^{N_{OUT}} w^{\dagger}_{kj}\left(\tau_{kp} + \varepsilon_k\right) \tag{B.6}$$

where x_{ip}, y_{jp} and z_{kp} are the neural activities for the neurons of the input, hidden and output layers (i, j and k respectively) for an input pattern (or stimulus) p. The targets of a pattern p for the hidden and output layers are y^*_{jp} and t_{kp}. The weight c_{ij} links the neurons x_i and y_j and the weight w_{jk} connects the neurons y_j and z_k. θ_j is the threshold of y_j and ε_k is the one for z_k. Finally, γ is the coefficient that defines the sigmoid slope.

Next, we are going to make the derivation of the four different terms of the equation with respect to the neural thresholds of θ and ε and the weights of W.

Derivative of the first term with respect to θ:

$$\begin{aligned}\frac{\partial E_1}{\partial \theta_n} &= \frac{\partial}{\partial \theta_j}\sum_{p=1}^{N_P}\alpha\left|y^*_p - y_p\right|^2 \\ &= \frac{\partial}{\partial \theta_n}\alpha\sum_{p=1}^{N_P}\left(y^*_p - y_p\right)^T\left(y^*_p - y_p\right) \\ &= \frac{\partial}{\partial \theta_n}\alpha\sum_{p=1}^{N_P}\left(y^{*T}_p y^*_p - y^{*T}_p y_p - y^T_p y^*_p + y^T_p y_p\right)\end{aligned} \tag{B.7}$$

Because $y*$ and y are vectors of the same dimensions:

$$y^{*T}_p y_p = y^T_p y^*_p = \sum_{j=1}^{N_{KC}} y^*_{jp} y_{jp} \tag{B.8}$$

$$y^{*T}_p y^*_p = \sum_{j=1}^{N_{KC}} y^{*^2}_{jp} \tag{B.9}$$

$$y^T_p y_p = \sum_{j=1}^{N_{KC}} y^2_{jp} \tag{B.10}$$

$$\frac{\partial E_1}{\partial \theta_n} = \frac{\partial}{\partial \theta_n}\alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} \left(y_{jp}^{*^2} - 2y_{jp}^{*}y_{jp} + y_{jp}^2 \right)$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} \left(\frac{\partial}{\partial \theta_n} y_{jp}^{*^2} - \frac{\partial}{\partial \theta_n} 2y_{jp}^{*}y_{jp} + \frac{\partial}{\partial \theta_n} y_{jp}^2 \right)$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} \left(-2y_{jp}^{*}\frac{\partial y_{jp}}{\partial \theta_n} + 2y_{jp}\frac{\partial y_{jp}}{\partial \theta_n} \right)$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp} - y_{jp}^{*}\right)\frac{\partial y_{jp}}{\partial \theta_n}$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp} - y_{jp}^{*}\right)\frac{\partial}{\partial \theta_n}\left[\frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}\right)} \right]$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp} - y_{jp}^{*}\right)\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}\right)^2}\frac{\partial}{\partial \theta_n}\left[exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)} \right] \tag{B.11}$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp} - y_{jp}^{*}\right)\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}$$

$$\frac{\partial}{\partial \theta_n}\left[-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j \right) \right]$$

$$= \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp} - y_{jp}^{*}\right)\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij}x_{ip} - \theta_j\right)}\gamma\delta_{n,j}$$

$$\delta_{n,j} = \begin{cases} 0 & n \neq j \\ 1 & n = j \end{cases}$$

$$= \alpha \sum_{p=1}^{N_P} 2\left(y_{np} - y_{np}^{*}\right)\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip} - \theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip} - \theta_n\right)}\gamma$$

Derivative of the first term with respect to ε:

$$
\begin{aligned}
\frac{\partial E_1}{\partial \varepsilon_l} &= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \left(\frac{\partial}{\partial \varepsilon_l} y_{jp}^{*^2} - \frac{\partial}{\partial \varepsilon_l} 2 y_{jp}^* y_{jp} + \frac{\partial}{\partial \varepsilon_l} y_{jp}^2 \right) \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \left(2 y_{jp}^* \frac{\partial y_{jp}^*}{\partial \varepsilon_l} - 2 y_{jp} \frac{\partial y_{jp}^*}{\partial \varepsilon_l} \right) \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \frac{\partial y_{jp}^*}{\partial \varepsilon_l} \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \frac{\partial}{\partial \varepsilon_l} \left[\sum_{k=1}^{N_{OUT}} w_{kj}^\dagger \left(\tau_{kp} + \varepsilon_k \right) \right] \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \sum_{k=1}^{N_{OUT}} w_{kj}^\dagger \delta_{l,k} \\
\delta_{l,k} &= \begin{cases} 0 & l \neq k \\ 1 & l = k \end{cases} \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) w_{lj}^\dagger
\end{aligned}
\tag{B.12}
$$

Derivative of the first term with respect to w:

$$
\begin{aligned}
\frac{\partial E_1}{\partial w_{nl}} &= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \left(\frac{\partial}{\partial w_{nl}} y_{jp}^{*^2} - \frac{\partial}{\partial w_{nl}} 2 y_{jp}^* y_{jp} + \frac{\partial}{\partial w_{nl}} y_{jp}^2 \right) \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \left(2 y_{jp}^* \frac{\partial y_{jp}^*}{\partial w_{nl}} - 2 y_{jp} \frac{\partial y_{jp}^*}{\partial w_{nl}} \right) \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \frac{\partial y_{jp}^*}{\partial w_{nl}} \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \frac{\partial}{\partial w_{nl}} \left[\sum_{k=1}^{N_{OUT}} w_{kj}^\dagger \left(\tau_{kp} + \varepsilon_k \right) \right] \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k \right) \frac{\partial w_{kj}^\dagger}{\partial w_{nl}} \\
&= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2 \left(y_{jp}^* - y_{jp} \right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k \right) \left(\frac{\partial w^\dagger}{\partial w_{nl}} \right)_{kj}
\end{aligned}
\tag{B.13}
$$

The pseudo-inverse is expressed as follows:

$$A^{\dagger} = \left(A^T A\right)^{-1} A^T \tag{B.14}$$

Then:

$$\begin{aligned}\frac{\partial E_1}{\partial w_{nl}} &= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2\left(y^*_{jp} - y_{jp}\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right) \frac{\partial}{\partial w_{nl}} \left[\left(w^T w\right)^{-1} w^T\right]_{kj} \\ &= \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2\left(y^*_{jp} - y_{jp}\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right) \left[\frac{\partial \left(w^T w\right)^{-1}}{\partial w_{nl}} w^T + \left(w^T w\right)^{-1} \frac{\partial w^T}{\partial w_{nl}}\right]_{kj}\end{aligned} \tag{B.15}$$

Being A a square matrix:

$$A^{-1}A = AA^{-1} = I \tag{B.16}$$

$$\frac{\partial I}{\partial A_{nl}} = 0 = \frac{\partial A^{-1} A}{\partial A_{nl}} \tag{B.17}$$

$$\frac{\partial A^{-1}}{\partial A_{nl}} A + A^{-1} \frac{\partial A}{\partial A_{nl}} = 0 \tag{B.18}$$

$$\frac{\partial A^{-1}}{\partial A_{nl}} A A^{-1} + A^{-1} \frac{\partial A}{\partial A_{nl}} A^{-1} = 0 \tag{B.19}$$

$$\frac{\partial A^{-1}}{\partial A_{nl}} = -A^{-1} \frac{\partial A}{\partial A_{nl}} A^{-1} \tag{B.20}$$

$\left(w^T w\right)^{-1}$ is a square matrix.

$$\begin{aligned}\frac{\partial E_1}{\partial w_{nl}} =& \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2\left(y^*_{jp} - y_{jp}\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right) \\ & \left[-\left(w^T w\right)^{-1} \frac{\partial \left(w^T w\right)}{\partial w_{nl}} \left(w^T w\right)^{-1} w^T + \left(w^T w\right)^{-1} I\left(n, l\right)\right]_{kj}\end{aligned} \tag{B.21}$$

$$I_{jk}\left(n, l\right) = \begin{cases} 1 & n = j, l = k \\ 0 & otherwise \end{cases} \tag{B.22}$$

$$\frac{\partial E_1}{\partial w_{nl}} = \alpha \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} 2\left(y^*_{jp} - y_{jp}\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right)$$

$$\left[\left(w^T w\right)^{-1}\left(-\frac{\partial\left(w^T w\right)}{\partial w_{nl}}\left(w^T w\right)^{-1} w^T + I\left(n,l\right)\right)\right]_{kj} \tag{B.23}$$

$$\begin{aligned}\left(w^T w\right)_{k_1 k_2} &= \sum_{r=1}^{N_{ROWS}} w_{k_1 r}^T w_{r k_2} = \sum_{r=1}^{N_{ROWS}} w_{r k_1} w_{r k_2} \\ &= \frac{\partial}{\partial w_{nl}}\left[\sum_{r=1}^{N_{ROWS}} w_{r k_1} w_{r k_2}\right] = \sum_{r=1}^{N_{ROWS}}\left[\frac{\partial w_{r k_1}}{\partial w_{nl}} w_{r k_2} + w_{r k_1}\frac{\partial w_{r k_2}}{\partial w_{nl}}\right] \\ &= \sum_{r=1}^{N_{ROWS}}\left[\delta_{r,n}\delta_{k_1,l} w_{r k_2} + w_{r k_1}\delta_{r,n}\delta_{k_2,l}\right] = \delta_{k_1,l} w_{n k_2} + w_{n k_1}\delta_{k,l}\end{aligned} \tag{B.24}$$

$$J_{k_1 k_2}\left(n,l\right) = \delta_{k_1,l} w_{n k_2} + w_{n k_1}\delta_{k_2,l} \tag{B.25}$$

$$\begin{aligned}\frac{\partial E_1}{\partial w_{nl}} =& \alpha \sum_{p=1}^{N_P}\sum_{j=1}^{N_{KC}} 2\left(y_{jp}^* - y_{jp}\right)\sum_{k=1}^{N_{OUT}}\left(\tau_{kp} + \varepsilon_k\right) \\ & \left[\left(w^T w\right)^{-1}\left(-J\left(n,l\right)\left(w^T w\right)^{-1} w^T + I\left(n,l\right)\right)\right]_{kj}\end{aligned} \tag{B.26}$$

Derivative of the second term with respect to θ:

$$\begin{aligned}\frac{\partial E_2}{\partial \theta_n} &= \frac{\partial}{\partial \theta_j}\sum_{p=1}^{N_P}\beta\left|t_p - z_p\right|^2 \\ &= \frac{\partial}{\partial \theta_n}\beta\sum_{p=1}^{N_P}\left(t_p - z_p\right)^T\left(t_p - z_p\right) \\ &= \frac{\partial}{\partial \theta_n}\beta\sum_{p=1}^{N_P}\left(t_p^T t_p - t_p^T z_p - z_p^T t_p + z_p^T z_p\right)\end{aligned} \tag{B.27}$$

Because t and z are vectors of the same dimensions:

$$\begin{aligned}\frac{\partial E_2}{\partial \theta_n} &= \frac{\partial}{\partial \theta_n}\beta\sum_{p=1}^{N_P}\sum_{k=1}^{N_{OUT}}\left(t_{kp}^2 - 2 t_{kp} z_{kp} + z_{kp}^2\right) \\ &= \beta\sum_{p=1}^{N_P}\sum_{k=1}^{N_{OUT}}\left(\frac{\partial}{\partial \theta_n} t_{kp}^2 - \frac{\partial}{\partial \theta_n} 2 t_{kp} z_{kp} + \frac{\partial}{\partial \theta_n} z_{kp}^2\right)\end{aligned} \tag{B.28}$$

Because z depends on y and this one depends on θ:

$$\begin{aligned}
\frac{\partial E_2}{\partial \theta_n} &= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} \left(-2t_{kp} \frac{\partial z_{kp}}{\partial \theta_n} + 2z_{kp} \frac{\partial z_{kp}}{\partial \theta_n} \right) \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{\partial z_{kp}}{\partial \theta_n} \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{\partial}{\partial \theta_n} \left[\frac{1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)} \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2} \frac{\partial}{\partial \theta_n} \left[exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} \right]
\end{aligned} \tag{B.29}$$

$$\begin{aligned}
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2} \\
&\quad exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} \frac{\partial}{\partial \theta_n} \left[-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right) \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2} \\
&\quad exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} - \gamma w_{jk} \frac{\partial y_{jp}}{\partial \theta_n}
\end{aligned}$$

$\frac{\partial y_j^p}{\partial \theta_n}$ was shown previously for the first term, therefore we can replace it here:

$$\begin{aligned}
\frac{\partial E_2}{\partial \theta_n} =& \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2} \\
& exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} - \gamma w_{jk} \\
& \left[\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)} \gamma \delta_{n,j} \right]
\end{aligned}$$

$$\delta_{n,j} = \begin{cases} 0 & n \neq j \\ 1 & n = j \end{cases} \tag{B.30}$$

$$= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2}$$

$$exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} w_{nk} \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_j\right)} \gamma^2$$

$$\frac{\partial E_2}{\partial \theta_n} = \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2}$$

$$exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} - \gamma w_{jk}$$

$$\left[\frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{ij} x_{ip} - \theta_j\right)} \gamma \delta_{n,j} \right]$$

$$\delta_{n,j} = \begin{cases} 0 & n \neq j \\ 1 & n = j \end{cases} \tag{B.31}$$

$$= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2\left(z_{kp} - t_{kp}\right) \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)}\right)^2}$$

$$exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k\right)} w_{nk} \frac{-1}{\left(1 + \exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_j\right)} \gamma^2$$

Derivative of the second term with respect to ε:

$$
\begin{aligned}
\frac{\partial E_2}{\partial \varepsilon_l} &= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} \left(\frac{\partial}{\partial \varepsilon_l} t_{kp}^2 - \frac{\partial}{\partial \varepsilon_l} 2 t_{kp} z_{kp} + \frac{\partial}{\partial \varepsilon_l} z_{kp}^2 \right) \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} \left(-2 t_{kp} \frac{\partial z_{kp}}{\partial \varepsilon_l} + 2 z_{kp} \frac{\partial z_{kp}}{\partial \varepsilon_l} \right) \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{\partial z_{kp}}{\partial \varepsilon_l} \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{\partial}{\partial \varepsilon_l} \left[\frac{1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)} \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad \frac{\partial}{\partial \varepsilon_l} \left[exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \frac{\partial}{\partial \varepsilon_l} \left[-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right) \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \gamma \delta_{l,k} \\
&\quad \delta_{l,k} = \begin{cases} 0 & l \neq k \\ 1 & l = k \end{cases} \\
&= \beta \sum_{p=1}^{N_P} 2 \left(t_{lp} - z_{lp} \right) \frac{1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)} \gamma
\end{aligned}
\tag{B.32}
$$

Derivative of the second term with respect to w:

$$
\begin{aligned}
\frac{\partial E_2}{\partial w_{nl}} &= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} \left(\frac{\partial}{\partial w_{nl}} t_{kp}^2 - \frac{\partial}{\partial w_{nl}} 2 t_{kp} z_{kp} + \frac{\partial}{\partial w_{nl}} z_{kp}^2 \right) \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} \left(-2 t_{kp} \frac{\partial z_{kp}}{\partial w_{nl}} + 2 z_{kp} \frac{\partial z_{kp}}{\partial w_{nl}} \right) \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{\partial z_{kp}}{\partial w_{nl}} \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{\partial}{\partial w_{nl}} \left[\frac{1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)} \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad \frac{\partial}{\partial w_{nl}} \left[exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \frac{\partial}{\partial w_{nl}} \left[-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right) \right] \\
&= \beta \sum_{p=1}^{N_P} \sum_{k=1}^{N_{OUT}} 2 \left(z_{kp} - t_{kp} \right) \frac{-1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)} - \gamma y_{jp} \delta_{n,j} \delta_{l,k} \\
&\quad \delta_{n,j} = \begin{cases} 0 & n \neq j \\ 1 & n = j \end{cases} \\
&\quad \delta_{l,k} = \begin{cases} 0 & l \neq k \\ 1 & l = k \end{cases} \\
&= \beta \sum_{p=1}^{N_P} 2 \left(z_{lp} - t_{lp} \right) \frac{1}{\left(1 + \exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)} \right)^2} \\
&\quad exp^{-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)} \gamma y_{np}
\end{aligned}
\tag{B.33}
$$

Derivative of the third term with respect to ε:

$$\frac{\partial E_3}{\partial \varepsilon_l} = \lambda \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \frac{\partial y_{jp}^*}{\partial \varepsilon_l} \tag{B.34}$$

$\frac{\partial y_j^{*p}}{\partial \varepsilon_l}$ was shown also for the first term, therefore we can replace it here:

$$\frac{\partial E_3}{\partial \varepsilon_l} = \lambda \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} w_{lj}^\dagger = \lambda N_P \sum_{j=1}^{N_{KC}} w_{lj}^\dagger \tag{B.35}$$

Derivative of the third term with respect to w:

$$\frac{\partial E_3}{\partial w_{nl}} = \frac{\partial}{\partial w_{nl}} \sum_{p=1}^{N_P} \lambda \left| y_p^* \right| = \lambda \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \frac{\partial}{\partial w_{nl}} \left| y_{jp}^* \right| = \lambda \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \frac{\partial y_{jp}^*}{\partial w_{nl}} \tag{B.36}$$

$\frac{\partial y_j^{*p}}{\partial w_{nl}}$ was shown previously for the first term, therefore we can replace it here:

$$\frac{\partial E_3}{\partial w_{nl}} = \lambda \sum_{p=1}^{N_P} \sum_{j=1}^{N_{KC}} \sum_{k=1}^{N_{OUT}} (\tau_{kp} + \varepsilon_k) \left[\left(w^T w \right)^{-1} \left(-J\left(n,l\right) \left(w^T w \right)^{-1} w^T + I\left(n,l\right) \right) \right]_{kj} \tag{B.37}$$

Derivative of the fourth term with respect to w:

$$\begin{aligned}
\frac{\partial E_4}{\partial w_{nl}} &= \frac{\partial}{\partial w_{nl}} \mu \frac{|w|^2}{2} = \frac{\partial}{\partial w_{nl}} \frac{\mu}{2} \sum_{j=1}^{N_{KC}} \sum_{k=1}^{N_{OUT}} w_{jk}^2 \\
&= \frac{\mu}{2} \sum_{j=1}^{N_{KC}} \sum_{k=1}^{N_{OUT}} \frac{\partial w_{jk}^2}{\partial w_{nl}} \\
&= \frac{\mu}{2} \sum_{j=1}^{N_{KC}} \sum_{k=1}^{N_{OUT}} 2 w_{jk} \delta_{n,j} \delta_{l,k} \qquad \text{(B.38)} \\
\delta_{n,j} &= \begin{cases} 0 & n \neq j \\ 1 & n = j \end{cases} \\
\delta_{l,k} &= \begin{cases} 0 & l \neq k \\ 1 & l = k \end{cases} \\
&= \frac{\mu}{2} 2 w_{nl} = \mu w_{nl}
\end{aligned}$$

Therefore, the derivatives of the equation B.2.1 with respect to the neural thresholds θ and ε and the weights w are:

$$\theta_n = \theta_n - \eta_\theta \frac{\partial E}{\partial \theta_n} \tag{B.39}$$

$$\begin{aligned}\frac{\partial E}{\partial \theta_n} = \sum_{p=1}^{N_P} \Bigg[& 2\alpha \left(y^*_{np} - y_{np}\right) \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\gamma \\ & + 2\beta \sum_{k=1}^{N_{OUT}} \left(t_{kp} - z_{kp}\right) \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk}y_{jp}-\varepsilon_k\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jk}y_{jp}-\varepsilon_k\right)} w_{nk} \\ & \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\right)^2} exp^{-\gamma\left(\sum_{i=1}^{N_{AL}} c_{in}x_{ip}-\theta_n\right)}\gamma^2 \Bigg]\end{aligned} \tag{B.40}$$

$$\varepsilon_l = \varepsilon_l - \eta_\varepsilon \frac{\partial E}{\partial \varepsilon_l} \tag{B.41}$$

$$\begin{aligned}\frac{\partial E}{\partial \varepsilon_l} = \sum_{p=1}^{N_P} \Bigg[& 2\alpha \sum_{j=1}^{N_{KC}} \left(y^*_{jp} - y_{jp}\right) w^\dagger_{lj} + 2\beta \left(t_{lp} - z_{lp}\right) \\ & \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\gamma \Bigg] + \lambda N_P \sum_{j=1}^{N_{KC}} w^\dagger_{lj}\end{aligned} \tag{B.42}$$

$$w_{nl} = w_{nl} - \eta_w \frac{\partial E}{\partial w_{nl}} \tag{B.43}$$

$$\begin{aligned}\frac{\partial E}{\partial w_{nl}} = \sum_{p=1}^{N_P} \Bigg[& \sum_{j=1}^{N_{KC}} \left(2\alpha \left(y_{jp} * -y_{jp}\right) + \lambda\right) \sum_{k=1}^{N_{OUT}} \left(\tau_{kp} + \varepsilon_k\right) \\ & \left[\left(w^T w\right)^{-1} \left(-J\left(n,l\right)\left(w^T w\right)^{-1} w^T + I\left(n,l\right)\right)\right]_{kj} + \beta \sum_{p=1}^{N_P} 2\left(z_{lp} - t_{lp}\right) \\ & \frac{1}{\left(1+\exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\right)^2} exp^{-\gamma\left(\sum_{j=1}^{N_{KC}} w_{jl}y_{jp}-\varepsilon_l\right)}\gamma y_{np} \Bigg] + \mu w_{nl}\end{aligned} \tag{B.44}$$

B.2.2 Final approach

Finally, as we mentioned in section 3.2.1.2, the equation B.2.1 had several problems. A high computational cost, difficulty in escaping from local minima and difficulty in balancing the different learning terms. Therefore, we reduced it to the following expression:

$$E = \alpha \left| T - Z \right|^2 + \beta \left| Y \right| \tag{B.45}$$

where T is the target output matrix, Z is the output matrix and Y is the hidden layer matrix. Both goals are weighted according to their coefficients, α for the classification error and β for the sparse neural activity.

The derivatives for these two terms with respect to the neural thresholds θ and ε and the weights of W in this case are:

$$\theta_n = \theta_n - \eta_\theta \frac{\partial E}{\partial \theta_n} \tag{B.46}$$

$$\frac{\partial E}{\partial \theta_n} = \sum_{p=1}^{N_P} \left[2\alpha \sum_{k=1}^{N_{MBON}} \left((t_{kp} - z_{kp}) \frac{\exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)) w_{nk}}{\left(1 + \exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jk} y_{jp} - \varepsilon_k \right)) \right)^2} \right. \right.$$
$$\left. \left. \frac{\exp(-\gamma \left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n \right)) \gamma^2}{\left(1 + \exp(-\gamma \left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n \right)) \right)^2} \right) - \beta \frac{\gamma \exp(-\gamma \left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n \right))}{\left(1 + \exp(-\gamma \left(\sum_{i=1}^{N_{AL}} c_{in} x_{ip} - \theta_n \right)) \right)^2} \right] \tag{B.47}$$

$$\varepsilon_l = \varepsilon_l - \eta_\varepsilon \frac{\partial E}{\partial \varepsilon_l} \tag{B.48}$$

$$\frac{\partial E}{\varepsilon_l} = 2\alpha \sum_{p=1}^{N_P} (t_{lp} - z_{lp}) \frac{\exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right))}{\left(1 + \exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)) \right)^2} \gamma \tag{B.49}$$

$$w_{nl} = w_{nl} - \eta_w \frac{\partial E}{\partial w_{nl}} \tag{B.50}$$

$$\frac{\partial E}{\partial w_{nl}} = 2\alpha \sum_{p=1}^{N_P} (z_{lp} - t_{lp}) \frac{\exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right))}{\left(1 + \exp(-\gamma \left(\sum_{j=1}^{N_{KC}} w_{jl} y_{jp} - \varepsilon_l \right)) \right)^2} \gamma y_{np} \tag{B.51}$$

APPENDIX

C

Complexity Measure

Using different types of patterns implies classifying data with different complexity. We assume that the complexity of a kind of pattern can have effects on the behavior of the neural network, so it would be useful to have a quantitative measure of complexity. Therefore, we decided to establish a complexity measure focused on pattern features [79], such as overlapping. This measure is based on the same fundamentals of Fisher discriminant [45, 79], Mahalanobis distance [127] and Silhouette [186], which are used as measures of class separation.

To measure the overlap, we calculate the intersection between patterns as follows:

$$O_{ij} = \frac{Area_i \cap Area_j}{Area_i \cup Area_j} \tag{C.1}$$

where O_{ij} is the intersection between the areas of two different patterns divided by their union. $Area_i$ and $Area_j$ are the representations in the two-dimensional space of the activity in PNs generated by the input patterns i and j (as we can see in Fig.3.11, where X-axis represents the PNs and Y-axis their activity).

This value is used to calculate the intraclass and interclass overlap. Being the intraclass overlap equations:

C. APPENDIX: COMPLEXITY MEASURE

$$O_k^{intra} = \frac{\sum_{i=1}^{N_{pc}-1} \sum_{j=i+1}^{N_{pc}} O_{ij}}{\binom{N_{pc}}{2}} \tag{C.2}$$

$$O^{intra} = \frac{\sum_{k=1}^{N_{class}} O_k^{intra}}{N_{class}} \tag{C.3}$$

where O_k^{intra} is the overlap between patterns from the class k. This is the overlap sum between all patterns, excluding repeated ones, divided by the number of their combinations. Where N_{pc} is the number of patterns for each class. Therefore, O^{intra} is the total intraclass overlap, which is calculated by the sum of each O_k^{intra} and divided by the number of classes, N_{class}.

This procedure changes for the interclass overlap, which is calculated as follows:

$$O_{kl}^{inter} = \frac{\sum_{i=1}^{N_{pc}} \sum_{j=1}^{N_{pc}} O_{ij}}{N_{pc}^2} \tag{C.4}$$

$$O^{inter} = \frac{\sum_{k=1}^{N_{class}-1} \sum_{l=k+1}^{N_{class}} O_{kl}^{inter}}{\binom{N_{class}}{2}} \tag{C.5}$$

where O_{kl}^{inter} is the overlap between patterns from two different classes (k and l) and O^{inter} their sum, excluding repeated ones, divided by the number of their combinations.

O^{inter}	O^{intra}	CL	Example in Fig. C.1
0	0	100	panel (a)
0	100	0	panel (b)
33	0	100	panel (c)
100	100	100	panel (d)
$O^{inter} \geq O^{intra}, CL = 100$ $O^{inter} < O^{intra}, CL < 100$			

Table C.1: Table of relationship between overlaps and complexity. It shows the relationship between overlaps (O^{inter} and O^{intra}) and the complexity level (CL) for the extreme cases for overlaps of Fig. C.1.

Once intra and interclass overlap, normalized by the number of classes and patterns, were defined (O^{intra} and O^{inter}), we developed a qualitative measure of

complexity that links both of them. The first thing that we had to do for this purpose was to study the relationship of both overlapping measures for their extreme values (see Fig.C.1 and Table.C.1). If $O^{inter} = 0\%$ and $O^{intra} = 100\%$, we will be in an ideal situation for pattern recognition because all patterns of the same class are identical and patterns of different classes are completely different, which implies a complexity of 0%. Otherwise, if $O^{inter} = 0\%$ and $O^{intra} = 0\%$ or $O^{inter} = 100\%$ and $O^{intra} = 100\%$, we will not be able to identify the class to which a pattern belongs. In fact, if $O^{inter}_{kl} = O^{intra}_{k} k \in \{1 \dots N_{class}\}$, the patterns are always indistinguishable and the complexity is 100%. Hence, we can generalize that when $O^{inter} = O^{intra}$, the complexity is total. Finally, if we have $O^{inter}_{kl} > O^{intra}_{k} k \in \{1 \dots N_{class}\}$, we cannot correctly classify and complexity takes its maximum value. Therefore, we can conclude that the complexity is 100% when $O^{inter} \geq O^{intra}$.

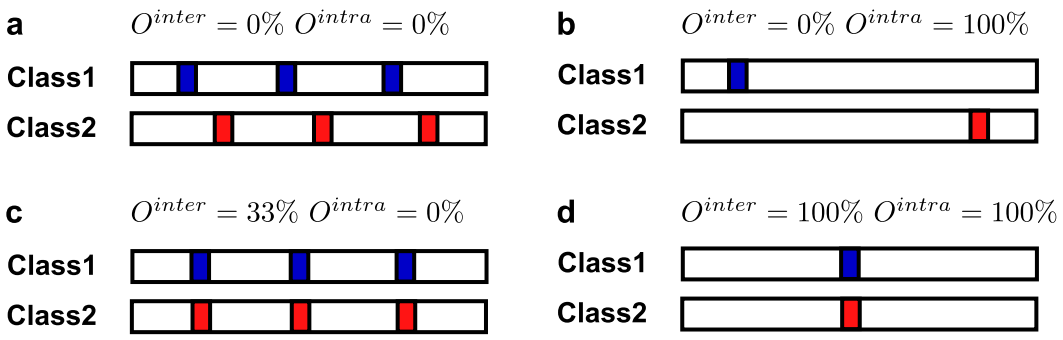


Figure C.1: Examples of intra and interclass overlap in extreme cases by using 2 pattern classes and 3 patterns of 1 dimension per class. In the case of $O^{inter} = 0\%$ and $O^{intra} = 0\%$, panel (a), the patterns of the same class do not share information between them, so although this information does not overlap with the patterns of other classes, in the worst case it will be impossible to achieve a separation between classes. In the case of $O^{inter} = 0\%$ and $O^{intra} = 100\%$, panel (b), all patterns for the same class are identical and patterns for different classes are completely different. In the case of $O^{inter} = 33\%$ and $O^{intra} = 0\%$ (because O^{inter} cannot reach 100% due to its maximum value depend on O^{intra}) both classes have identical patterns that do not share information between them and, therefore, they only overlap with one third of the other class's patterns, panel (c). This makes it impossible to discern which pattern belongs to each class. In the case of $O^{inter} = 100\%$ and $O^{intra} = 100\%$, panel (d), all patterns are the same pattern and, therefore, they are impossible to separate.

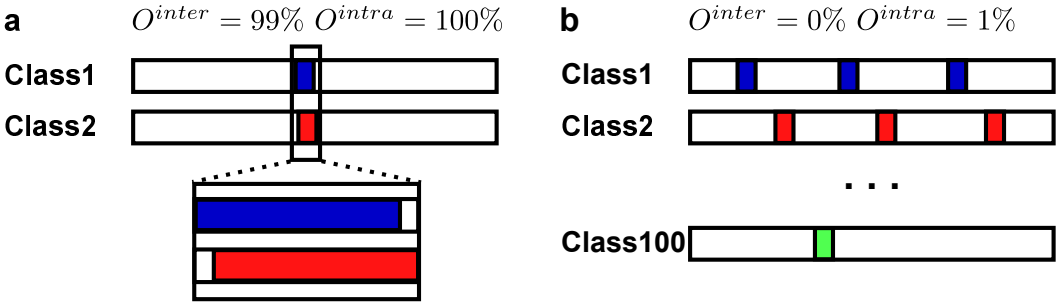


Figure C.2: Examples of $O^{inter} < O^{intra}$ in cases close to their extremes for 3 patterns per class of 1 dimension. In the case of $O^{inter} = 99\%$ and $O^{intra} = 100\%$, panel (a), the patterns of each class are identical and the slightest difference between classes is sufficient to distinguish them. In the case of $O^{inter} = 0\%$, $O^{intra} = 1\%$, panel (b), we assume the worst case for pattern recognition, where 99 classes have $O^{intra} = 0\%$ and one class has $O^{intra} = 100\%$, $Class100$. This last class, whose patterns are identical, is the only one that can be correctly classified.

However, it remains to know what happens when $O^{inter} < O^{intra}$ and the complexity is not 100% (see Fig. C.2 and Table. C.2). To understand the relationship between complexity and overlaps for this condition, we studied two cases close to the extremes. In the first one, we established $O^{intra} = 100\%$ and gave an arbitrary value to O^{inter}, for example $O^{inter} = 99\%$. If $O^{inter}_{kl} < O^{intra}_{k} k \in \{1 \dots N_{class}\}$, we can ensure that even if the value of O^{inter} is high, as in our case, there is no difficulty in classifying and therefore the complexity is 0. This is possible because the slightest difference between the pattern classes will be sufficient to distinguish their patterns, since the patterns of a class are identical. In the second case, we established $O^{inter} = 0\%$ and gave another value to O^{intra}, for example $O^{intra} = 1\%$. Because $O^{inter} = 0\%$ and $O^{intra} \neq 0\%$, the complexity should be 0. However, since O^{intra} is the average value of all pattern classes, in the worst case, we would have that 1% of patterns are identical to patterns of their own class, $O^{intra} = 100\%$, while 99% of them share no information with the rest, $O^{intra} = 0\%$. Therefore, classification success in the worst case is equal to O^{intra}, since we can only classify 1% of patterns. This implies a complexity of $100\% - O^{intra}$. The reason for choosing this relationship between O^{intra} and complexity is that it allows us to punish cases where classes are not well-defined. Although it is true that for $O^{inter} = 0\%$ the complexity must tend quickly to zero, once $O^{inter} \neq 0\%$, its complexity will

be severely affected by the value of O^{intra}. Although this evolution does not have to be linear, as we assume here, this assumption allows us to make an acceptable approximation as shown by the empirical data of Fig. C.5.

O^{inter}	O^{intra}	CL	Example in Fig. C.2
X	100	0	panel (a)
0	X	100-X	panel (b)

Table C.2: Table of relationship between overlaps and complexity when $O^{inter} < O^{intra}$. It shows the relationship between overlaps (O^{inter} and O^{intra}) and the complexity level (CL) when $O^{inter} < O^{intra}$. If $O^{intra} = 100\%$ and we have a value X for O^{inter} that is $X < O^{intra}$, the complexity is 0. If $O^{inter} = 0\%$ and we have a value X for O^{intra} that is $X > O^{inter}$, the complexity is $100\% - O^{intra}$.

Keeping this in mind, we define the following equation for complexity:

$$CL = \frac{100 - O^{intra}}{100 - O^{inter}} \times 100 \tag{C.6}$$

where CL is the complexity level when $O^{inter} < O^{intra}$. If O^{intra} is low, the complexity will be high. Otherwise, if O^{intra} is high, the complexity will be low, until this reaches 0 for $O^{intra} = 100$ and independently of O^{inter} value. On the other hand, O^{inter} behaves inversely to O^{intra}. Therefore, when $O^{inter} = 0$, the complexity percentage will be $100 - O^{intra}$, Fig. C.3. Since the overlap measures, O^{intra} and O^{inter}, are normalized based on the number of classes and patterns per class, the complexity measure, CL, is too. The fact that the complexity measure is normalized allows limiting its value between 0 and 100, since complexity is generally an unbounded measure [181, 220]. However, this normalization involves that the complexity introduced by the number of classes or patterns per class used cannot be calculated. Therefore, the results for electronic-nose data (6 classes instead of 10) in Figs. C.5 and 4.11 should be shifted slightly to the left, although this displacement does not affect the conclusions we have drawn from these figures.

In Fig. C.4, we can observe the complexity and overlaps (inter and intraclass) for all input patterns that we used. The overlap is from $\sim 0\%$ to 62.93% for Gaussian patterns, 21.5% for MNIST digits and 61.37% for electronic-nose data, between patterns of different classes (interclass). Otherwise, between patterns of

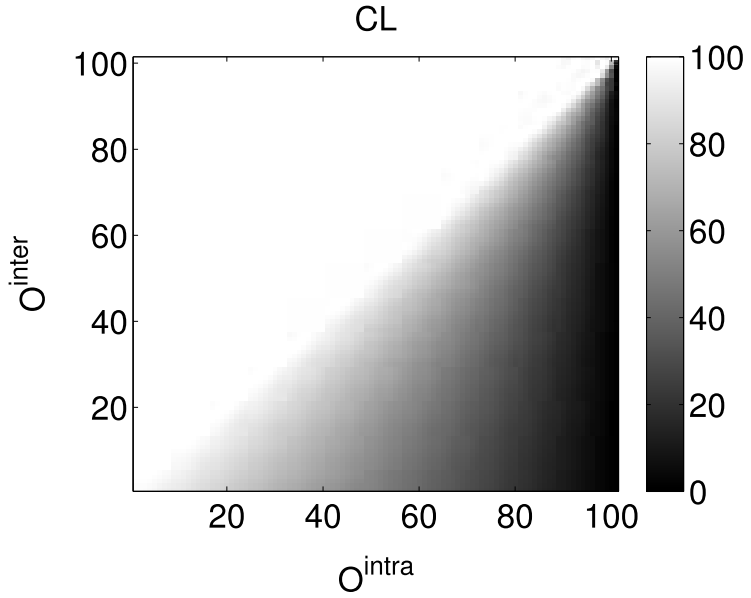


Figure C.3: Relationship between overlaps and complexity. Complexity values (CL) for different combination of overlaps (O^{inter} and O^{intra}). When $O^{inter} < O^{intra}$, the complexity is calculated by Eq. C.6. Otherwise, the complexity is 100%, as it was deduced from the extreme cases, Table.C.1.

the same class (intraclass), the overlap is from 70.69% to 92.13% for Gaussian patterns, 32.62% for MNIST digits and 80.98% for electronic-nose data. This low intraclass overlap between MNIST digits can explain their high complexity, 85.83%, despite their having a smaller interclass overlap than electronic-nose inputs, which have a complexity of 49.23%. This complexity for MNIST digits is also greater than the one of the Gaussian patterns that we showed in Fig. 3.11, from 7.87% to 77.64%, but not for all Gaussian patterns that we used (see Figs. C.5 and 4.11) with STD values from 1 to 200.

Finally, we compare our complexity measure, CL, with the classification errors for different classifiers and 5-cross-validation, Fig. C.5. Since we used Gaussian patterns to study the effects of different overlapping degrees (complexity), we can use these patterns as a reference. For them, the relationship between classification error and the proposed complexity measure is directly proportional. For LDA, Naive Bayes and linear SVM classifiers, the error increases slightly exponentially as the complexity augments. In the case of KNN classifier, this growth of the error is

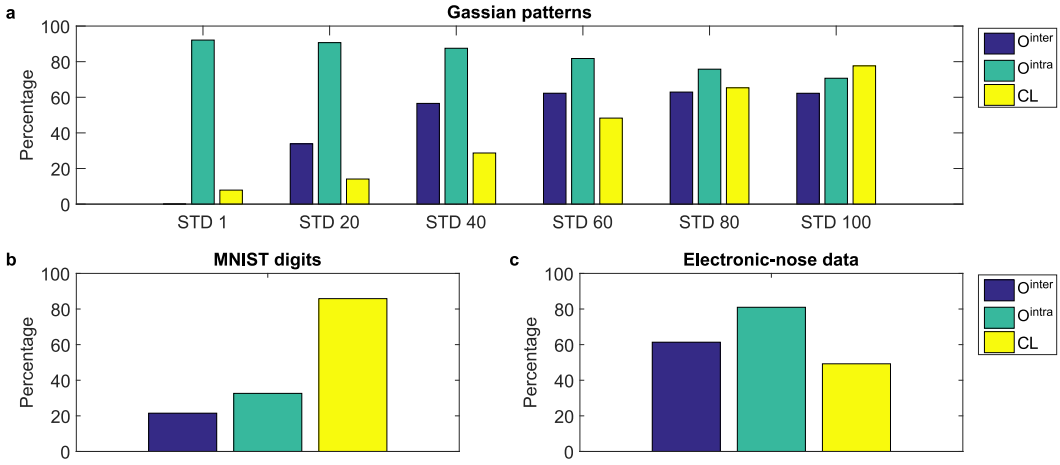


Figure C.4: Overlaps and complexity in patterns. a: this bar graph shows the overlapping and complexity values for Gaussian patterns. The interclass overlap (O^{inter}) go from $\sim 0\%$ to 62.93%. In case of, intraclass overlap (O^{intra}), their values are from 70.69% to 92.13%. Finally, the complexity (CL) of the Gaussian patterns (for STD values from 1 to 100) go from 7.87% to 77.64%. b: this figure shows the overlapping values for the MNIST, 21.5% and 32.62% (inter and intraclass), and their complexity, 85.83%. c: this panel shows these values for electronic-nose data, where the overlapping values are 61.37% and 80.98% (inter and intraclass) and the complexity is 49.23%.

more abrupt, similar to a sigmoid, and the one in which our method works worse. On the other hand, we can observe that electronic-nose data (eNOSE), whose complexity should be reduced because of its smaller number of classes, matches with the relationship between classification error and complexity of Gaussian patterns. This does not happen for MNIST digits, where the classification error is significantly lower than the one obtained by Gaussian patterns, expect for Naive Bayes classifier, for the same complexity. The most likely explanation for this behavior is due to our complexity equation punishes the fact that classes are not well-defined (low O^{intra}). Therefore, in the case of the MNIST digits with an $O^{intra} = 32.62\%$, they obtain a greater complexity than was recorded by some classifiers. Although there are other differences between these datasets, such as the use of binary or float data and the number of neurons involved in coding. However, we can say that our complexity equation is able to qualitatively measure the complexity of patterns,

which obtains very similar values to those we would get with a Naive Bayes classifier.

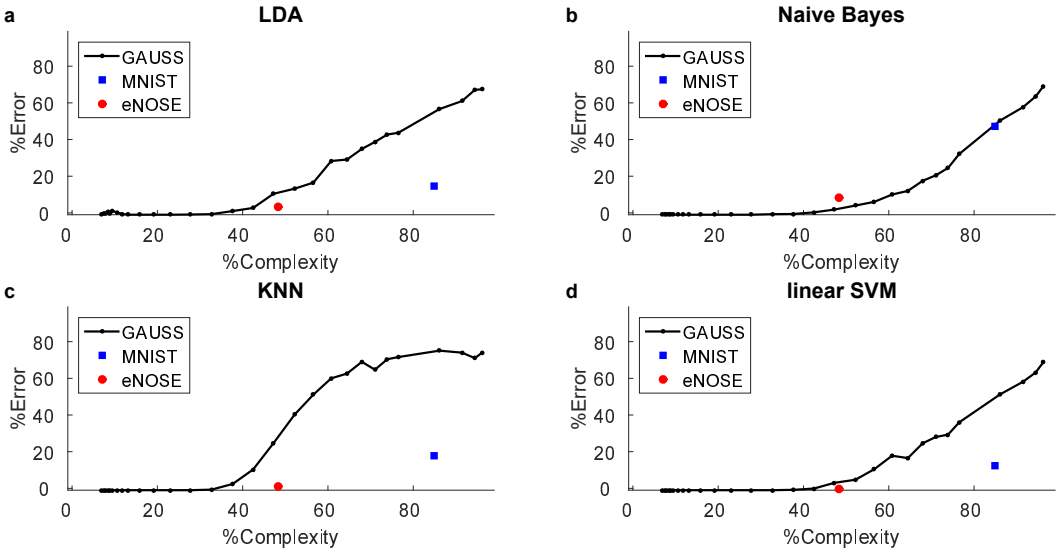


Figure C.5: Relationship between complexity and classification error for different classifiers. These figures show how the classification error increases when the complexity of Gaussian patterns, calculated by the Eq. C.6, also augments. When we compare these results with those obtained with electronic-nose (eNOSE) data and MNIST digits, it is observed that the first ones match for all classifiers while the last ones only do it for Naive Bayes. The most likely explanation for this behavior is because of our complexity equation punishes the fact that classes are not well-defined (low O^{intra}). Therefore, the MNIST digits, $O^{intra} = 32.62\%$, obtain a greater complexity than the one estimated by some classifiers. Although there are also other differences between these datasets, such as the use of binary or float data and the number of neurons involved in coding. Nevertheless, these results show that our complexity equation is able to qualitatively measure the complexity of patterns.

Bibliography

[1] AlphaPsi. Schematic insect basiconic sensillum. https://commons.wikimedia.org/wiki/File:Schematic_insect_basiconic_sensillum.png. CC BY-SA 4.0. 17

[2] Anna Maria Angioy, Alessandro Desogus, Iole Tomassini Barbarossa, Peter Anderson, and Bill S Hansson. Extreme sensitivity in an olfactory system. *Chem Senses*, 28(4):279–284, May 2003. 8, 36, 113

[3] Sylvia Anton and Uwe Homberg. Antennal lobe structure. In *Insect olfaction*, pages 97–124. Springer, 1999. 19, 50

[4] Sylvia Anton, Rickard Ignell, and Bill S Hansson. Developmental changes in the structure and function of the central olfactory system in gregarious and solitary desert locusts. *Microscopy research and technique*, 56(4):281–291, 2002. 19

[5] J Douglas Armstrong, Kim Kaiser, Armin Müller, Karl-Friedrich Fischbach, Nirav Merchant, and Nicholas J Strausfeld. Flybrain, an on-line atlas and database of the drosophila nervous system. *Neuron*, 15(1):17–20, 1995. 6, 61, 91, 97

[6] Collins Assisi, Mark Stopfer, Gilles Laurent, and Maxim Bazhenov. Adaptive regulation of sparseness by feedforward inhibition. *Nat Neurosci*, 10(9):1176–1184, Sep 2007. 21

[7] O Ávila-Åkerberg, R Krahe, and MJ Chacron. Neural heterogeneities and stimulus properties affect burst coding in vivo. *Neuroscience*, 168(1):300–313, 2010. 8, 24, 36

[8] NJ Bannister and AU Larkman. Dendritic morphology of ca1 pyramidal neurones from the rat hippocampus: I. branching patterns. *Journal of Comparative Neurology*, 360(1):150–160, 1995. 24

[9] Brice Bathellier, Samuel Lagier, Philippe Faure, and Pierre-Marie Lledo. Circuit properties generating gamma oscillations in a network model of the olfactory bulb. *Journal of neurophysiology*, 95(4):2678–2691, 2006. 22

[10] Maxim Bazhenov, Ramon Huerta, and Brian H Smith. A computational framework for understanding decision making through integration of basic learning rules. *The Journal of Neuroscience*, 33(13):5686–5697, 2013. 32, 33

[11] Maxim Bazhenov, Mark Stopfer, Mikhail Rabinovich, Henry DI Abarbanel, Terrence J Sejnowski, and Gilles Laurent. Model of cellular and network mechanisms for odor-evoked temporal patterning in the locust antennal lobe. *Neuron*, 30(2):569–581, 2001. 22, 23

[12] Gerd Bicker, Sabine Schäfer, and Timothy G Kingan. Mushroom body feedback interneurones in the honeybee show gaba-like immunoreactivity. *Brain research*, 360(1-2):394–397, 1985. 8, 22

[13] Eric Block, Seogjoo Jang, Hiroaki Matsunami, Sivakumar Sekharan, Bérénice Dethier, Mehmed Z Ertem, Sivaji Gundala, Yi Pan, Shengju Li, Zhen Li, et al. Implausibility of the vibrational theory of olfaction. *Proceedings of the National Academy of Sciences*, 112(21):E2766–E2774, 2015. 14

[14] J Boeckh and Leslie P Tolbert. Synaptic organization and development of the antennal lobe in insects. *Microscopy research and technique*, 24(3):260–280, 1993. 18

[15] Thomas Bresson. Caelifera sp. https://www.flickr.com/photos/computerhotline/3871567734. CC BY 2.0. 17

[16] Romain Brette. Computing with neural synchrony. *PLoS computational biology*, 8(6):e1002561, 2012. 24

[17] Linda Buck and Richard Axel. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*, 65(1):175–187, 1991. 16

[18] Caroline Bushdid, Marcelo O Magnasco, Leslie B Vosshall, and Andreas Keller. Humans can discriminate more than 1 trillion olfactory stimuli. *Science*, 343(6177):1370–1372, 2014. 5, 91, 97

[19] John R Carlson. Functional expression of a drosophila odor receptor. *Proceedings of the National Academy of Sciences*, 98(16):8936–8937, 2001. 16

[20] Mikael A Carlsson, C Giovanni Galizia, and Bill S Hansson. Spatial representation of odours in the antennal lobe of the moth spodoptera littoralis (lepidoptera: Noctuidae). *Chemical senses*, 27(3):231–244, 2002. 6, 17, 18

[21] Sophie JC Caron, Vanessa Ruta, LF Abbott, and Richard Axel. Random convergence of olfactory inputs in the drosophila mushroom body. *Nature*, 497(7447):113, 2013. 7, 20, 31

[22] Stijn Cassenaer and Gilles Laurent. Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature*, 448(7154):709–713, Aug 2007. 8, 16, 27, 44

[23] Stijn Cassenaer and Gilles Laurent. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, 482(7383):47, 2012. 8, 16, 27

[24] Ho Ka Chan and Thomas Nowotny. A biophysical model of the early olfactory system of honeybees. In *International Conference on Neural Information Processing*, pages 639–647. Springer, 2017. 23

[25] Gal Chechik, Isaac Meilijson, and Eytan Ruppin. Neuronal regulation: A mechanism for synaptic pruning during brain maturation. *Neural computation*, 11(8):2061–2080, 1999. 45

[26] Mircea I Chelaru and Valentin Dragoi. Efficient coding in heterogeneous neuronal populations. *Proceedings of the National Academy of Sciences*, 2008. 24

[27] A Chess, L Buck, MM Dowling, R Axel, and J Ngai. Molecular biology of smell: expression of the multigene family encoding putative odorant receptors. 57:505–516, 1992. 16

[28] Thomas A Christensen. Making scents out of spatial and temporal codes in specialist and generalist olfactory networks. *Chemical senses*, 30(suppl_-1):i283–i284, 2005. 8, 9, 24, 25, 43, 44, 61, 69, 91, 92, 98, 99

[29] François David, Christiane Linster, and Thomas A Cleland. Lateral dendritic shunt inhibition can regularize mitral cell spike patterning. *Journal of computational neuroscience*, 25(1):25–38, 2008. 22

[30] Ronald L Davis. Olfactory learning. *Neuron*, 44(1):31–48, 2004. 19

[31] Andrew P Davison, Jianfeng Feng, and David Brown. Dendrodendritic inhibition and simulated odor responses in a detailed olfactory bulb network model. *Journal of neurophysiology*, 90(3):1921–1935, 2003. 22

[32] J Steven de Belle and Martin Heisenberg. Associative odor learning in drosophila abolished by chemical ablation of mushroom bodies. *Science*, 263(5147):692–695, 1994. 22

[33] Claire A de March, SangEun Ryu, Gilles Sicard, Cheil Moon, and Jérôme Golebiowski. Structure–odour relationships reviewed in the postgenomic era. *Flavour and Fragrance Journal*, 30(5):342–361, 2015. 12

[34] Michael Denker, Marc Timme, Markus Diesmann, Fred Wolf, and Theo Geisel. Breaking synchrony by heterogeneity in complex networks. *Physical review letters*, 92(7):074103, 2004. 24

[35] Brett J Doleman and Nathan S Lewis. Comparison of odor detection thresholds and odor discriminablities of a conducting polymer composite electronic nose versus mammalian olfaction. *Sensors and Actuators B: Chemical*, 72(1):41–50, 2001. 37, 92, 95, 98, 101

[36] Josh Dubnau, Lori Grady, Toshi Kitamoto, and Tim Tully. Disruption of neurotransmission in drosophila mushroom body blocks retrieval but not acquisition of memory. *Nature*, 411(6836):476, 2001. 31, 33

[37] Katharina Eichler, Feng Li, Ashok Litwin-Kumar, Youngser Park, Ingrid Andrade, Casey M Schneider-Mizell, Timo Saumweber, Annina Huser, Claire Eschbach, Bertram Gerber, et al. The complete connectome of a learning and memory centre in an insect brain. *Nature*, 548(7666):175, 2017. 7, 22, 83, 84, 85

[38] KD Ernst and J Boeckh. A neuroanatomical study on the organization of the central antennal pathways in insects. *Cell and tissue research*, 229(1):1–22, 1983. 19, 50

[39] KD Ernst, J Boeckh, and V Boeckh. A neuroanatomical study on the organization of the central antennal pathways in insects. *Cell and tissue research*, 176(3):285–308, 1977. 21

[40] Monty A Escabí, Reza Nassiri, Lee M Miller, Christoph E Schreiner, and Heather L Read. The contribution of spike threshold to acoustic feature selectivity, spike information content, and information throughput. *Journal of Neuroscience*, 25(41):9524–9534, 2005. 25, 37

[41] Susan E Fahrbach. Structure of the mushroom bodies of the insect brain. *Annu. Rev. Entomol.*, 51:209–232, 2006. 20

[42] Sarah M Farris. Evolution of insect mushroom bodies: old clues, new insights. *Arthropod structure & development*, 34(3):211–234, 2005. 19, 20

[43] Sarah M Farris and Nicholas J Strausfeld. Development of laminar organization in the mushroom bodies of the cockroach: Kenyon cell proliferation, outgrowth, and maturation. *Journal of Comparative Neurology*, 439(3):331–351, 2001. 6, 61, 91, 97

[44] André Fiala, Thomas Spall, Sören Diegelmann, Beate Eisermann, Silke Sachse, Jean-Marc Devaud, Erich Buchner, and C.Giovanni Galizia. Genetically expressed cameleon in drosophila melanogaster is used to visualize

olfactory information in projection neurons. *Current Biology*, 12(21):1877 – 1884, 2002. 6, 17, 18

[45] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936. 133

[46] Daniel Flanagan and Alison R Mercer. An atlas and 3-d reconstruction of the antennal lobes in the worker honey bee, apis mellifera l.(hymenoptera: Apidae). *International Journal of Insect Morphology and Embryology*, 18(2-3):145–159, 1989. 19, 50

[47] Jordi Fonollosa, Alexander Vergara, and Ramón Huerta. Algorithmic mitigation of sensor failure: Is sensor replacement really necessary? *Sensors and Actuators B: Chemical*, 2013. 37, 92, 95, 98, 101

[48] Bertrand Fontaine, José Luis Peña, and Romain Brette. Spike-threshold adaptation predicted by membrane potential dynamics in vivo. *PLoS computational biology*, 10(4):e1003560, 2014. 24, 25, 37

[49] Nicolas Fourcaud-Trocmé, David Hansel, Carl Van Vreeswijk, and Nicolas Brunel. How spike generation mechanisms determine the neuronal response to fluctuating inputs. *Journal of Neuroscience*, 23(37):11628–11640, 2003. 25, 37

[50] Maria Isabel Franco, Luca Turin, Andreas Mershin, and Efthimios MC Skoulakis. Molecular vibration-sensing component in drosophila melanogaster olfaction. *Proceedings of the National Academy of Sciences*, 108(9):3797–3802, 2011. 14

[51] Marion E Frank. Neuron types, receptors, behavior, and taste quality. *Physiology & behavior*, 69(1-2):53–62, 2000. 24, 25, 44

[52] Marion E Frank, Stephen L Bieber, and David V Smith. The organization of taste sensibilities in hamster chorda tympani nerve fibers. *The Journal of general physiology*, 91(6):861–896, 1988. 24, 25, 44

[53] Hans U Fried, Stefan H Fuss, and Sigrun I Korsching. Selective imaging of presynaptic activity in the mouse olfactory bulb shows concentration and structure dependence of odor responses in identified glomeruli. *Proceedings of the National Academy of Sciences*, 99(5):3222–3227, 2002. 14, 28, 54

[54] C Giovanni Galizia and Randolf Menzel. Probing the olfactory code. *nature neuroscience*, 3(9):853, 2000. 61, 91, 97

[55] C Giovanni Galizia, Silke Sachse, and Hanna Mustaparta. Calcium responses to pheromones and plant odours in the antennal lobe of the male and female moth heliothis virescens. *Journal of Comparative Physiology A*, 186(11):1049–1063, 2000. 6, 17, 18

[56] C Giovanni Galizia, Silke Sachse, Angelika Rappert, and Randolf Menzel. The glomerular code for odor representation is species specific in the honeybee apis mellifera. *Nature neuroscience*, 2(5):473, 1999. 14

[57] C.Giovanni Galizia and Randolf Menzel. The role of glomeruli in the neural representation of odours: results from optical recording studies. *Journal of Insect Physiology*, 47(2):115 – 130, 2001. 6, 17, 18

[58] Cosmas Giovanni Galizia and B Kimmerle. Physiological and morphological characterization of honeybee olfactory neurons combining electrophysiology, calcium imaging and confocal microscopy. *Journal Of Comparative Physiology A*, 190(1):21–38, 2004. 6, 61, 91, 97

[59] Cosmas Giovanni Galizia, Sabrina L McIlwrath, and Randolf Menzel. A digital three-dimensional atlas of the honeybee antennal lobe based on optical sections acquired by confocal microscopy. *Cell and tissue research*, 295(3):383–394, 1999. 6, 61, 91, 97

[60] Cosmas Giovanni Galizia, Randolf Menzel, and Bert Hölldobler. Optical imaging of odor-evoked glomerular activity patterns in the antennal lobes of the ant camponotus rufipes. *Naturwissenschaften*, 86(11):533–537, 1999. 6, 17, 18

[61] Simon Gane, Dimitris Georganakis, Klio Maniati, Manolis Vamvakias, Nikitas Ragoussis, Efthimios MC Skoulakis, and Luca Turin. Molecular vibration-sensing component in human olfaction. *PloS one*, 8(1):e55780, 2013. 14

[62] Marta García-Sanchez and Ramón Huerta. Design parameters of the fan-out phase of sensory systems. *Journal of computational neuroscience*, 15(1):5–17, 2003. 7, 21, 31, 63, 92, 98

[63] J Gascuel and C Masson. A quantitative ultrastructural study of the honeybee antennal lobe. *Tissue and Cell*, 23(3):341–355, 1991. 18

[64] Michael Gazzaniga and Richard B Ivry. *Cognitive Neuroscience: The Biology of the Mind: Fourth International Student Edition*. WW Norton, 2013. 5

[65] Cyrille Claude Girardin, Sabine Kreissl, and Cosmas Giovanni Galizia. Inhibitory connections in the honeybee antennal lobe are spatially patchy. *American Journal of Physiology-Heart and Circulatory Physiology*, 2012. 7, 25, 50, 61, 92, 94, 98, 100

[66] D Golomb, D Hansel, and G Mato. Mechanisms of synchrony of neural activity in large networks. In *Handbook of biological physics*, volume 4, pages 887–968. Elsevier, 2001. 24

[67] David Golomb and John Rinzel. Dynamics of globally coupled inhibitory neurons with heterogeneity. *Physical review E*, 48(6):4810, 1993. 24

[68] Mario González, David Dominguez, Guillermo Jerez, and Odette Pantoja. Periodically diluted begnn model of corruption perception. In *International Conference on Mining Intelligence and Knowledge Exploration*, pages 289–298. Springer, 2018. 95, 101

[69] Mario González, David Dominguez, Francisco B Rodríguez, and Angel Sanchez. Retrieval of noisy fingerprint patterns using metric attractor networks. *International journal of neural systems*, 24(07):1450025, 2014. 95, 101

[70] Mario González, David Dominguez, Ángel Sánchez, and Francisco B Rodríguez. Capacity and retrieval of a modular set of diluted attractor networks with respect to the global number of neurons. In *International Work-Conference on Artificial Neural Networks*, pages 497–506. Springer, 2017. 95, 101

[71] Mario González, David Dominguez, Ángel Sánchez, and Francisco B Rodríguez. Increase attractor capacity using an ensembled neural network. *Expert Systems with Applications*, 71:206–215, 2017. 95, 101

[72] Nitin Gupta and Mark Stopfer. Functional analysis of a higher olfactory center, the lateral horn. *Journal of Neuroscience*, 32(24):8138–8148, 2012. 22

[73] Michael Häusser and Bartlett Mel. Dendrites: bug or feature? *Current opinion in neurobiology*, 13(3):372–383, 2003. 24

[74] Donald Olding Hebb. *The Organization of Behavior*. Wiley, 1949. 33

[75] Gertrud Heimbeck, Véronique Bugnon, Nanaë Gendre, Andreas Keller, and Reinhard F Stocker. A central neural circuit for experience-independent olfactory and courtship behavior in drosophila melanogaster. *Proceedings of the National Academy of Sciences*, 98(26):15336–15341, 2001. 22

[76] Martin Heisenberg. Mushroom body memoir: from maps to models. *Nat Rev Neurosci*, 4(4):266–275, Apr 2003. 19

[77] DA Henze and G Buzsaki. Action potential threshold of hippocampal pyramidal cells in vivo is increased by recent spiking activity. *Neuroscience*, 105(1):121–130, 2001. 24, 25, 37, 92, 98

[78] John G Hildebrand and Gordon M Shepherd. Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annual review of neuroscience*, 20(1):595–631, 1997. 6, 61, 91, 97

[79] Tin Kam Ho and Mitra Basu. Complexity measures of supervised classification problems. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (3):289–300, 2002. 133

[80] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500–544, 1952. 22

[81] Lars A Holmstrom, Lonneke BM Eeuwes, Patrick D Roberts, and Christine V Portfors. Efficient encoding of vocalizations in the auditory midbrain. *Journal of Neuroscience*, 30(3):802–819, 2010. 8, 36

[82] Uwe Homberg, Thomas A Christensen, and JG Hildebrand. Structure and function of the deutocerebrum in insects. *Annual review of entomology*, 34(1):477–501, 1989. 21

[83] Uwe Homberg, Timothy G Kingan, and John G Hildebrand. Immunocytochemistry of gaba in the brain and suboesophageal ganglion ofmanduca sexta. *Cell and tissue research*, 248(1):1–24, 1987. 22

[84] Elizabeth J Hong and Rachel I Wilson. Simultaneous encoding of odors by channels with diverse sensitivity to inhibition. *Neuron*, 85(3):573–589, 2015. 7

[85] MacKenzie A Howard and Edwin W Rubel. Dynamic spike thresholds during synaptic integration preserve and enhance temporal response properties in the avian cochlear nucleus. *Journal of Neuroscience*, 30(36):12063–12074, 2010. 24, 25, 37

[86] Chao Huang, Andrey Resnik, Tansu Celikel, and Bernhard Englitz. Adaptive spike threshold enables robust and temporally precise neuronal encoding. *PLoS computational biology*, 12(6):e1004984, 2016. 24, 25, 37, 92, 98

[87] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3):489–501, 2006. 33, 36, 41, 120

[88] Ramon Huerta, Thiago Mosqueiro, Jordi Fonollosa, Nikolai F Rulkov, and Irene Rodriguez-Lujan. Online decorrelation of humidity and temperature in chemical sensors for continuous monitoring. *Chemometrics and Intelligent Laboratory Systems*, 157:169–176, 2016. 94, 101

[89] Ramon Huerta and Thomas Nowotny. Fast and robust learning by reinforcement signals: Explorations in the insect brain. *Neural Comput.*, 21:2123–2151, 2009. 12, 22, 23, 27, 28, 31, 34, 59, 61, 75, 91, 97

[90] Ramon Huerta, Thomas Nowotny, Marta García-Sanchez, Henry DI Abarbanel, and Mikhail I Rabinovich. Learning classification in the olfactory system of insects. *Neural computation*, 16(8):1601–1640, 2004. 12, 22, 23, 28, 31, 33, 73

[91] Eric Hunsberger, Matthew Scott, and Chris Eliasmith. The competing benefits of noise and heterogeneity in neural coding. *Neural computation*, 26(8):1600–1623, 2014. 24

[92] Rickard Ignell, Sylvia Anton, and Bill S Hansson. The antennal lobe of orthoptera–anatomy and evolution. *Brain, behavior and evolution*, 57(1):1–17, 2001. 19

[93] Martin Cooper Ipswich. Fruit fly (drosophila immigrans). `https://www.flickr.com/photos/92899351@N08/13114869053`. CC BY 2.0. 15

[94] Kei Ito, Kazunori Shinomiya, Masayoshi Ito, J Douglas Armstrong, George Boyan, Volker Hartenstein, Steffen Harzsch, Martin Heisenberg, Uwe Homberg, Arnim Jenett, et al. A systematic nomenclature for the insect brain. *Neuron*, 81(4):755–765, 2014. 6

[95] Kei Ito, Kazumi Suzuki, Patricia Estes, Mani Ramaswami, Daisuke Yamamoto, and Nicholas J Strausfeld. The organization of extrinsic neurons and their implications in the functional roles of the mushroom bodies in drosophila melanogaster meigen. *Learning & Memory*, 5(1):52–77, 1998. 6, 61, 91, 97

[96] Vladimir Itskov, Carina Curto, Eva Pastalkova, and György Buzsáki. Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus. *Journal of Neuroscience*, 31(8):2828–2834, 2011. 24, 25, 37

[97] USCDCP Janice Carr, Connie Flowers. https://pixnio.com/science/microscopy-images/insects/segmentation-along-its-length-allowed-for-greater-flexibility-of-the-antennae. CC 0. 17

[98] Gregory SXE Jefferis, Christopher J Potter, Alexander M Chan, Elizabeth C Marin, Torsten Rohlfing, Calvin R Maurer, and Liqun Luo. Comprehensive maps of drosophila higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell*, 128(6):1187–1203, 2007. 22

[99] Shozo Jinno, Thomas Klausberger, Laszlo F Marton, Yannis Dalezios, J David B Roberts, Pablo Fuentealba, Eric A Bushong, Darrell Henze, György Buzsáki, and Peter Somogyi. Neuronal diversity in gabaergic long-range projections from the hippocampus. *Journal of Neuroscience*, 27(33):8790–8804, 2007. 24

[100] Douglas L Jones, Erik C Johnson, and Rama Ratnam. A stimulus-dependent spike threshold is an optimal neural coder. *Frontiers in computational neuroscience*, 9:61, 2015. 25, 37

[101] Ron A Jortner, S. Sarah Farivar, and Gilles Laurent. A simple connectivity scheme for sparse coding in an olfactory system. *J Neurosci*, 27(7):1659–1669, Feb 2007. 7, 8, 63

[102] U Benjamin Kaupp. Olfactory signalling in vertebrates and insects: differences and commonalities. *Nature Reviews Neuroscience*, 11(3):188, 2010. 8, 9, 25, 43, 69, 92, 99

[103] Tiffany Kee, Pavel Sanda, Nitin Gupta, Mark Stopfer, and Maxim Bazhenov. Feed-forward versus feedback inhibition in a basic olfactory circuit. *PLoS computational biology*, 11(10):e1004531, 2015. 14

[104] Andreas Keller and Leslie B Vosshall. A psychophysical test of the vibration theory of olfaction. *Nature neuroscience*, 7(4):337, 2004. 14

[105] FC Kenyon. The brain of the bee. a preliminary contribution to the morphology of the nervous system of the arthropoda. *Journal of Comparative Neurology*, 6(3):133–210, 1896. 19

[106] Asami Kido and Kei Ito. Mushroom bodies are not required for courtship behavior by normal and sexually mosaic drosophila. *Developmental Neurobiology*, 52(4):302–311, 2002. 22

[107] Jan Kropf and Wolfgang Rössler. In-situ recording of ionic currents in projection neurons and kenyon cells in the olfactory pathway of the honeybee. *PloS one*, 13(1):e0191425, 2018. 8

[108] Samuel Lagier, Patrizia Panzanelli, Raúl E Russo, Antoine Nissant, Brice Bathellier, Marco Sassoè-Pognetto, Jean-Marc Fritschy, and Pierre-Marie Lledo. Gabaergic inhibition at dendrodendritic synapses tunes γ oscillations in the olfactory bulb. *Proceedings of the National Academy of Sciences*, 104(17):7259–7264, 2007. 22

[109] Gilles Laurent. A systems perspective on early olfactory coding. *Science*, 286(5440):723–728, 1999. 7

[110] Gilles Laurent. Olfactory network dynamics and the coding of multidimensional signals. *Nature reviews neuroscience*, 3(11):884, 2002. 8, 15

[111] Gilles Laurent and Mohammad Naraghi. Odorant-induced oscillations in the mushroom bodies of the locust. *Journal of Neuroscience*, 14(5):2993–3004, 1994. 44

[112] Gilles Laurent, Mark Stopfer, Rainer W Friedrich, Misha I Rabinovich, Alexander Volkovskii, and Henry DI Abarbanel. Odor encoding as an active, dynamical process: experiments, computation, and theory. *Annual review of neuroscience*, 24(1):263–297, 2001. 15

[113] Gilles Laurent, Michael Wehr, and Hananel Davidowitz. Temporal representations of odors in an olfactory network. *Journal of Neuroscience*, 16(12):3837–3847, 1996. 19

[114] Walter S Leal. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annual review of entomology*, 58:373–391, 2013. 16

[115] Yann LeCun, Corinna Cortes, and Christopher JC Burges. The mnist database of handwritten digits, 1998, 1998. 28, 54, 59, 73

[116] Beulah Leitch and Gilles Laurent. Gabaergic synapses in the antennal lobe and mushroom body of the locust olfactory system. *Journal of comparative Neurology*, 372(4):487–514, 1996. 19, 44, 50

[117] Andrew C Lin, Alexei M Bygrave, Alix De Calignon, Tzumin Lee, and Gero Miesenböck. Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. *Nature neuroscience*, 17(4):559, 2014. 8

[118] Tao Lin, Chaofeng Li, Jiali Liu, Brian H Smith, Hong Lei, and Xinnian Zeng. Glomerular organization in the antennal lobe of the oriental fruit fly bactrocera dorsalis. *Frontiers in neuroanatomy*, 12, 2018. 17, 18

[119] Christiane Linster and Brian H Smith. A computational model of the response of honey bee antennal lobe circuitry to odor mixtures: overshadowing, blocking and unblocking can arise from lateral inhibition. *Behavioural brain research*, 87(1):1–14, 1997. 23

[120] Xu Liu and Ronald L Davis. The gabaergic anterior paired lateral neuron suppresses and is suppressed by olfactory learning. *Nature neuroscience*, 12(1):53, 2009. 22

[121] Fernando F Locatelli, Patricia C Fernandez, and Brian H Smith. Learning about natural variation of odor mixtures enhances categorization in early olfactory processing. *Journal of Experimental Biology*, 219(17):2752–2762, 2016. 5

[122] Jessica Lopez-Hazas. Desarrollo de una toolbox para medir la respuesta positiva de un sistema hacia estímulos externos. B.S. thesis, Universidad Autónoma de Madrid, Madrid, Spain, 2017. 13, 81, 82, 83, 93, 100

[123] Jessica Lopez-Hazas. Propiedades computacionales del umbral neuronal y el código sparse en el proceso de discriminaciónde odorantes. Master's thesis, Universidad Autónoma de Madrid, Madrid, Spain, 2018. 36, 43, 62, 68, 69, 114

[124] Jessica Lopez-Hazas, Aaron Montero, and Francisco B Rodriguez. Strategies to enhance pattern recognition in neural networks based on the insect olfactory system. In *International Conference on Artificial Neural Networks*, pages 468–475. Springer, 2018. 43, 62, 68, 69, 103, 114

[125] Stefano Luccioli and Antonio Politi. Irregular collective behavior of heterogeneous neural networks. *Physical review letters*, 105(15):158104, 2010. 24

[126] Robert F Lundy Jr and Robert J Contreras. Gustatory neuron types in rat geniculate ganglion. *Journal of neurophysiology*, 82(6):2970–2988, 1999. 25, 44

[127] P.C. Mahalanobis. On the generalised distance in statistics. *Proceedings of the National Institute of Science of India*, 2(1):49–55, 1936. 133

[128] Troy W Margrie and Andreas T Schaefer. Theta oscillation coupled spike latencies yield computational vigour in a mammalian sensory system. *The Journal of physiology*, 546(2):363–374, 2003. 22

[129] Elizabeth C Marin, Gregory SXE Jefferis, Takaki Komiyama, Haitao Zhu, and Liqun Luo. Representation of the glomerular olfactory map in the drosophila brain. *Cell*, 109(2):243–255, 2002. 20, 22

[130] Gary Marsat and Leonard Maler. Neural heterogeneity and efficient population codes for communication signals. *Journal of neurophysiology*, 104(5):2543–2555, 2010. 8, 36

[131] Liria M Masuda-Nakagawa, Kei Ito, Takeshi Awasaki, and Cahir J O'Kane. A single gabaergic neuron mediates feedback of odor-evoked signals in the mushroom body of larval drosophila. *Frontiers in neural circuits*, 8:35, 2014. 8

[132] Liria M Masuda-Nakagawa, Nobuaki K Tanaka, and Cahir J O'Kane. Stereotypic and random patterns of connectivity in the larval mushroom body calyx of drosophila. *Proceedings of the National Academy of Sciences*, 102(52):19027–19032, 2005. 7, 20, 31

[133] SG Matsumoto and JG Hildebrand. Olfactory mechanisms in the moth manduca sexta: response characteristics and morphology of central neurons in the antennal lobes. *Proceedings of the Royal Society of London B: Biological Sciences*, 213(1192):249–277, 1981. 19, 50

[134] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943. 22

[135] Sean E McGuire, Phuong T Le, and Ronald L Davis. The role of drosophila mushroom body signaling in olfactory memory. *Science*, 293(5533):1330–1333, 2001. 31

[136] JF Mejias and A Longtin. Optimal heterogeneity for coding in spiking neural networks. *Physical Review Letters*, 108(22):228102, 2012. 24

[137] Jorge F Mejias and André Longtin. Differential effects of excitatory and inhibitory heterogeneity on the gain and asynchronous state of sparse cortical networks. *Frontiers in computational neuroscience*, 8:107, 2014. 7, 24, 25, 50, 94, 100

[138] Anna Menini. *The neurobiology of olfaction*. CRC Press, 2009. 16

[139] Michele Migliore and Gordon M Shepherd. Dendritic action potentials connect distributed dendrodendritic microcircuits. *Journal of computational neuroscience*, 24(2):207–221, 2008. 22

[140] Makoto Mizunami, Josette M Weibrecht, and Nicholas J Strausfeld. Mushroom bodies of the cockroach: their participation in place memory. *Journal of Comparative Neurology*, 402(4):520–537, 1998. 6, 61, 91, 97

[141] Yamile Molina and Sean O'Donnell. Mushroom body volume is related to social aggression and ovary development in the paperwasp polistes instabilis. *Brain, behavior and evolution*, 70(2):137–144, 2007. 19, 20

[142] Robert Wighton Moncrieff. What is odor? a new theory. *Am. Perfumer*, 54(45):3, 1949. 14

[143] P Read Montague, Peter Dayan, Christophe Person, and Terrence J Sejnowski. Bee foraging in uncertain environments using predictive hebbian learning. *Nature*, 377(6551):725, 1995. 33

[144] Aaron Montero, Ramon Huerta, and Francisco B Rodriguez. Neuron threshold variability in an olfactory model improves odorant discrimination. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 16–25. Springer, 2013. 28, 31, 32, 33, 38, 39, 48, 54, 55, 59, 62, 63, 73, 86, 115

[145] Aaron Montero, Ramon Huerta, and Francisco B Rodriguez. Neural tradeoffs among specialist and generalist neurons in pattern recognition. In *International Conference on Engineering Applications of Neural Networks*, pages 71–80. Springer, 2014. 30, 31, 32, 34, 38, 40, 46, 48, 58, 59, 62, 86

[146] Aaron Montero, Ramon Huerta, and Francisco B Rodriguez. Regulation of specialists and generalists by neural variability improves pattern recognition performance. *Neurocomputing*, 151:69–77, 2015. 28, 31, 32, 33, 38, 40, 48, 54, 55, 62, 64, 68, 73, 83, 84, 86, 116

[147] Aaron Montero, Ramon Huerta, and Francisco B Rodriguez. Specialist neurons in feature extraction are responsible for pattern recognition process in insect olfaction. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 58–67. Springer, 2015. 30, 31, 32, 34, 38, 40, 48, 59, 73, 75, 86

[148] Aaron Montero, Ramon Huerta, and Francisco B Rodriguez. Stimulus space complexity determines the ratio of specialist and generalist neurons during pattern recognition. *Journal of the Franklin Institute*, 355(5):2951–2977, 2018. 30, 31, 34, 38, 40, 48, 58, 59, 62, 76, 86

[149] Aaron Montero, Jessica Lopez-Hazas, and Francisco B Rodriguez. Input pattern complexity determines specialist and generalist populations in drosophila neural network. In *International Conference on Artificial Neural Networks*, pages 296–303. Springer, 2018. 30, 31, 34, 38, 40, 48, 58, 62, 86

[150] Aaron Montero, Thiago Mosqueiro, Ramon Huerta, and Francisco B. Rodriguez. *Exploring a Mathematical Model of Gain Control via Lateral Inhibition in the Antennal Lobe*, pages 317–326. Springer International Publishing, Cham, 2017. 30, 31, 34, 38, 40, 48, 58, 86

[151] Kensaku Mori and Gordon M Shepherd. Emerging principles of molecular signal processing by mitral/tufted cells in the olfactory bulb. In *Seminars in cell biology*, volume 5, pages 65–74. Elsevier, 1994. 14

[152] Thiago S Mosqueiro and Leonardo P Maia. Optimal channel efficiency in a sensory network. *Physical Review E*, 88(1):012712, 2013. 58

[153] Mala Murthy, Ila Fiete, and Gilles Laurent. Testing odor response stereotypy in the drosophila mushroom body. *Neuron*, 59(6):1009–1023, 2008. 7, 30

[154] Tatsuro Nakagawa, Maurizio Pellegrino, Koji Sato, Leslie B Vosshall, and Kazushige Touhara. Amino acid residues contributing to function of the heteromeric insect olfactory receptor complex. *PloS one*, 7(3):e32372, 2012. 16

[155] Takanori Nakane. Activatoin-adenylate cyclase-outlined. https://commons.wikimedia.org/wiki/File:Activatoin-Adenylate_cyclase-outlined.svg. CC BY-SA 3.0. 18

[156] L Neltner, David Hansel, Germán Mato, and Claude Meunier. Synchrony in heterogeneous networks of spiking neurons. *Neural computation*, 12(7):1607–1641, 2000. 24

[157] Minna Ng, Robert D Roorda, Susana Q Lima, Boris V Zemelman, Patrick Morcillo, and Gero Miesenböck. Transmission of olfactory information between three populations of neurons in the antennal lobe of the fly. *Neuron*, 36(3):463–474, 2002. 52

[158] Yoshihito Niimura, Atsushi Matsui, and Kazushige Touhara. Extreme expansion of the olfactory receptor gene repertoire in african elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome research*, 24(9):1485–1496, 2014. 5, 91, 97

[159] Srinath Nizampatnam, Debajit Saha, Rishabh Chandak, and Baranidharan Raman. Dynamic contrast enhancement and flexible odor codes. *Nature communications*, 9(1):3062, 2018. 7

[160] Thomas Nowotny. Parallel implementation of a spiking neuronal network model of unsupervised olfactory learning on nvidia® cuda™. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2010. 22, 23

[161] Thomas Nowotny, Ramon Huerta, Henry DI Abarbanel, and Mikhail I Rabinovich. Self-organization in the olfactory system: one shot odor recognition in insects. *Biological cybernetics*, 93(6):436–446, 2005. 21, 22, 23

[162] Thomas Nowotny, Mikhail I Rabinovich, Ramon Huerta, and Henry DI Abarbanel. Decoding temporal information through slow lateral excitation in the olfactory system of insects. *Journal of computational neuroscience*, 15(2):271–281, 2003. 22

[163] Thomas Nowotny, Jacob S Stierle, C Giovanni Galizia, and Paul Szyszka. Data-driven honeybee antennal lobe model suggests how stimulus-onset asynchrony can aid odour segregation. *Brain research*, 1536:119–134, 2013. 23

[164] Thomas Nowotny and Paul Szyszka. Dynamics of odor-evoked activity patterns in the olfactory system. In *Advances in Dynamics, Patterns, Cognition*, pages 243–261. Springer, 2017. 7

[165] Yuki Oka, Masayo Omura, Hiroshi Kataoka, and Kazushige Touhara. Olfactory receptor antagonism between odorants. *The EMBO journal*, 23(1):120–126, 2004. 13

[166] Simona Olmi, Roberto Livi, Antonio Politi, and Alessandro Torcini. Collective oscillations in disordered neural networks. *Physical Review E*, 81(4):046119, 2010. 24

[167] Shawn R Olsen, Vikas Bhandawat, and Rachel I Wilson. Divisive normalization in olfactory population codes. *Neuron*, 66(2):287–299, 2010. 7, 48, 49, 52, 53, 86, 94

[168] Shawn R Olsen and Rachel I Wilson. Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature*, 452(7190):956, 2008. 7, 19, 48, 49, 52, 53, 86, 88, 94, 100

[169] Randall C O'reilly. Generalization in interactive networks: The benefits of inhibitory competition and hebbian learning. *Neural computation*, 13(6):1199–1241, 2001. 8, 16, 21, 23

[170] Leslie C Osborne, Stephanie E Palmer, Stephen G Lisberger, and William Bialek. The neural basis for combinatorial coding in a cortical population response. *Journal of Neuroscience*, 28(50):13522–13531, 2008. 8, 36

[171] Swidbert R Ott, Andrew Philippides, Maurice R Elphick, and Michael O'Shea. Enhanced fidelity of diffusive nitric oxide signalling by the spatial segregation of source and target neurones in the memory centre of an insect brain. *European Journal of Neuroscience*, 25(1):181–190, 2007. 20

[172] Maria Papadopoulou, Stijn Cassenaer, Thomas Nowotny, and Gilles Laurent. Normalization for sparse encoding of odors by a wide-field interneuron. *Science*, 332(6030):721–725, May 2011. 8, 22

[173] Toni Pérez, Claudio R Mirasso, Raúl Toral, and James D Gunton. The constructive role of diversity in the global response of coupled neuron systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1933):5619–5632, 2010. 24

[174] Javier Perez-Orive, Maxim Bazhenov, and Gilles Laurent. Intrinsic and circuit properties favor coincidence detection for decoding oscillatory input. *J Neurosci*, 24(26):6037–6047, Jun 2004. 8, 24

[175] Javier Perez-Orive, Ofer Mazor, Glenn C Turner, Stijn Cassenaer, Rachel I Wilson, and Gilles Laurent. Oscillations and sparsening of odor representations in the mushroom body. *Science*, 297(5580):359–365, Jul 2002. 7, 8, 13, 21, 25, 30, 31, 34, 36, 39, 41, 44, 61, 66, 71, 81, 82, 83, 84, 91, 92, 93, 97, 98, 100, 113, 120

[176] Jena L Pitman, Wolf Huetteroth, Christopher J Burke, Michael J Krashes, Sen-Lin Lai, Tzumin Lee, and Scott Waddell. A pair of inhibitory neurons are required to sustain labile memory in the drosophila mushroom body. *Current Biology*, 21(10):855–861, 2011. 22

[177] Nicholas J Priebe, Ferenc Mechler, Matteo Carandini, and David Ferster. The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nature neuroscience*, 7(10):1113, 2004. 37

[178] MI Rabinovich, R Huerta, A Volkovskii, HDI Abarbanel, M Stopfer, and G Laurent. Dynamical coding of sensory information with competitive networks. *Journal of Physiology-Paris*, 94(5-6):465–471, 2000. 32

[179] Russell Reed. Pruning algorithms-a survey. *IEEE transactions on Neural Networks*, 4(5):740–747, 1993. 45

[180] Alex Reyes, Rafael Lujan, Andrej Rozov, Nail Burnashev, Peter Somogyi, and Bert Sakmann. Target-cell-specific facilitation and depression in neocortical circuits. *Nature neuroscience*, 1(4):279, 1998. 24

[181] Francisco B Rodríguez and Ramón Huerta. Analysis of perfect mappings of the stimuli through neural temporal sequences. *Neural networks*, 17(7):963–973, 2004. 81, 137

[182] Francisco B Rodríguez and Ramón Huerta. Techniques for temporal detection of neural sensitivity to external stimulation. *Biological cybernetics*, 100(4):289–297, 2009. 8, 24, 25, 44, 45, 61, 69, 81, 82, 91, 93, 98, 100, 117

[183] Francisco B Rodríguez, Ramón Huerta, and Maria de la Luz Aylwin. Neural sensitivity to odorants in deprived and normal olfactory bulbs. *PloS one*, 8(4):e60745, 2013. 24, 25, 44, 61, 69, 91, 98

[184] Irene Rodriguez-Lujan, Jordi Fonollosa, Alexander Vergara, Margie Homer, and Ramon Huerta. On the calibration of sensor arrays for pattern recognition using the minimal number of experiments. *Chemometrics and Intelligent Laboratory Systems*, 2013. 28, 54, 59, 73

[185] Hai-Jun Rong, Yew-Soon Ong, Ah-Hwee Tan, and Zexuan Zhu. A fast pruned-extreme learning machine for classification problem. *Neurocomputing*, 72(1-3):359–366, 2008. 33, 36

[186] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987. 133

[187] Benjamin D Rubin and Lawrence C Katz. Optical imaging of odorant representations in the mammalian olfactory bulb. *Neuron*, 23(3):499–511, 1999. 56, 58

[188] Vanessa Ruta, Sandeep Robert Datta, Maria Luisa Vasconcelos, Jessica Freeland, Loren L Looger, and Richard Axel. A dimorphic pheromone circuit in drosophila from sensory input to descending output. *Nature*, 468(7324):686, 2010. 22

[189] Majid Saberi and Hamed Seyed-Allaei. Odorant receptors of drosophila are sensitive to the molecular volume of odorants. *Scientific reports*, 6:25103, 2016. 14

[190] Silke Sachse and C. Giovanni Galizia. The coding of odour-intensity in the honeybee antennal lobe: local computation optimizes odour representation. *Eur J Neurosci*, 18(8):2119–2132, Oct 2003. 56, 58

[191] Emilio Salinas and Peter Thier. Gain modulation: a major computational principle of the central nervous system. *Neuron*, 27(1):15–21, 2000. 7, 48

[192] Pavel Sanda, Tiffany Kee, Nitin Gupta, Mark Stopfer, and Maxim Bazhenov. Classification of odorants across layers in locust olfactory pathway. *Journal of neurophysiology*, 115(5):2303–2316, 2016. 7, 14

[193] Jean-Christophe Sandoz. Behavioral and neurophysiological study of olfactory perception and learning in honeybees. *Frontiers in systems neuroscience*, 5:98, 2011. 7, 50, 61, 92, 94

[194] Harvey B Sarnat, Laura Flores-Sarnat, and Xing-Chang Wei. Olfactory development, part 1: Function, from fetal perception to adult wine-tasting. *Journal of child neurology*, 32(6):566–578, 2017. 5

[195] M Savard, R Krahe, and MJ Chacron. Neural heterogeneities influence envelope and temporal coding at the sensory periphery. *Neuroscience*, 172:270–284, 2011. 8, 24, 36

[196] Simone Scardapane and Dianhui Wang. Randomness in neural networks: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(2):e1200, 2017. 93, 95, 99, 101

[197] Michael Schmuker, Nobuhiro Yamagata, Martin Nawrot, and Randolf Menzel. Parallel representation of stimulus identity and intensity in a dual pathway model inspired by the olfactory system of the honeybee. *Frontiers in Neuroengineering*, 4(17), 2011. 23

[198] Janna N Schultzhaus, Sehresh Saleem, Hina Iftikhar, and Ginger E Carney. The role of the drosophila lateral horn in olfactory information processing and behavioral response. *Journal of insect physiology*, 98:29–37, 2017. 7

[199] F Schürmann, Ina Frambach, and K Elekes. Gabaergic synaptic connections in mushroom bodies of insect brains. *Acta Biologica Hungarica*, 59(Supplement 2):173–181, 2008. 8, 16, 27, 32

[200] Yoichi Seki, Hany KM Dweck, Jürgen Rybak, Dieter Wicher, Silke Sachse, and Bill S Hansson. Olfactory coding from the periphery to higher brain centers in the drosophila brain. *BMC biology*, 15(1):56, 2017. 7

[201] Eduardo Serrano, Thomas Nowotny, Rafael Levi, Brian H Smith, and Ramón Huerta. Gain control network conditions in early sensory coding. *PLoS computational biology*, 9(7), 2013. 7, 48

[202] Hung-Chang Shen, Jia-Yi Wei, Sao-Yu Chu, Pei-Chi Chung, Tsai-Chi Hsu, and Hung-Hsiang Yu. Morphogenetic studies of the drosophila da1 ventral olfactory projection neuron. *PloS one*, 11(5):e0155384, 2016. 30, 50

[203] Hanne Therese Skiri, Cosmas Giovanni Galizia, and Hanna Mustaparta. Representation of primary plant odorants in the antennal lobe of the moth heliothis virescens using calcium imaging. *Chemical Senses*, 29(3):253–267, 2004. 6, 17, 18

[204] Richard B Stein. Some models of neuronal variability. *Biophysical journal*, 7(1):37–68, 1967. 22

[205] Reinhard F Stocker, Gertrud Heimbeck, Nanae Gendre, and J Steven de Belle. Neuroblast ablation in drosophila p [gal4] lines reveals origins of olfactory interneurons. *Journal of neurobiology*, 32(5):443–456, 1997. 50, 52

[206] Mark Stopfer. Central processing in the mushroom bodies. *Current Opinion in Insect Science*, 6:99 – 103, 2014. Pests and resistance/Parasites/Parasitoids/Biological control/Neurosciences. 8, 15, 16

[207] Mark Stopfer, Vivek Jayaraman, and Gilles Laurent. Intensity versus identity coding in an olfactory system. *Neuron*, 39(6):991–1004, 2003. 7, 18, 48

[208] Martin Strauch, Mathias Ditzen, and Cosmas Giovanni Galizia. Keeping their distance? odor response patterns along the concentration range. *Frontiers in systems neuroscience*, 6:71, 2012. 14, 28, 54, 56

[209] Nicholas J Strausfeld. Organization of the honey bee mushroom body: representation of the calyx within the vertical and gamma lobes. *Journal of Comparative Neurology*, 450(1):4–33, 2002. 6, 61, 91, 97

[210] Nicholas J Strausfeld and Cole Gilbert. Small-field neurons associated with oculomotor control in muscoid flies: Cellular organization in the lobula plate. *Journal of Comparative Neurology*, 316(1):56–71, 1992. 19

[211] Nicholas J Strausfeld, Lars Hansen, Yongsheng Li, Robert S Gomez, and Kei Ito. Evolution, discovery, and interpretations of arthropod mushroom bodies. *Learning & memory*, 5(1):11–37, 1998. 19

[212] Xue-Jun Sun, Caroline Fonta, and Claudine Masson. Odour quality processing by bee antennal lobe interneurones. *Chemical Senses*, 18(4):355–377, 1993. 19, 50

[213] Paul Szyszka, Mathias Ditzen, Alexander Galkin, C Giovanni Galizia, and Randolf Menzel. Sparsening and temporal sharpening of olfactory representations in the honeybee mushroom bodies. *Journal of neurophysiology*, 94(5):3303–3313, 2005. 8

[214] Paul Szyszka, Mathias Ditzen, Alexander Galkin, C Giovanni Galizia, and Randolf Menzel. Sparsening and temporal sharpening of olfactory representations in the honeybee mushroom bodies. *Journal of neurophysiology*, 94(5):3303–3313, 2005. 66, 92, 98

[215] Sachin S Talathi, Dong-Uk Hwang, and William L Ditto. Spike timing dependent plasticity promotes synchrony of inhibitory networks in the presence of heterogeneity. *Journal of Computational Neuroscience*, 25(2):262–281, 2008. 24

[216] Sachin S Talathi, Dong-Uk Hwang, Abraham Miliotis, Paul R Carney, and William L Ditto. Predicting synchrony in heterogeneous pulse coupled oscillators. *Physical Review E*, 80(2):021908, 2009. 24

[217] Nobuaki K Tanaka, Takeshi Awasaki, Takashi Shimada, and Kei Ito. Integration of chemosensory pathways in the drosophila second-order olfactory centers. *Current biology*, 14(6):449–457, 2004. 22

[218] T Tateno and HPC Robinson. Rate coding and spike-time variability in cortical neurons with two types of threshold dynamics. *Journal of neurophysiology*, 95(4):2650–2663, 2006. 37

[219] Claudio J Tessone, Claudio R Mirasso, Raúl Toral, and James D Gunton. Diversity-induced resonance. *Physical review letters*, 97(19):194101, 2006. 24

[220] Giulio Tononi and Gerald M Edelman. Consciousness and complexity. *science*, 282(5395):1846–1851, 1998. 137

[221] Marco Trincavelli, Alexander Vergara, Nikolai Rulkov, Jose S Murguia, Achim Lilienthal, and Ramon Huerta. Optimizing the operating temperature for an array of mox sensors on an open sampling system. 1362(1):225–227, 2011. 94, 101

[222] Luca Turin. A spectroscopic mechanism for primary olfactory reception. *Chemical senses*, 21(6):773–791, 1996. 14

[223] Glenn C Turner, Maxim Bazhenov, and Gilles Laurent. Olfactory representations by drosophila mushroom body neurons. *Journal of neurophysiology*, 99(2):734–746, 2008. 7

[224] JJA van Loon and LM Schoonhoven. Specialist deterrent chemoreceptors enable pieris butterflies to discriminate chemically different deterrents. pages 16–16, 1998. 24, 25, 44

[225] Alexander Vergara, Shankar Vembu, Tuba Ayhan, Margaret A Ryan, Margie L Homer, and Ramón Huerta. Chemical gas sensor drift compensation using classifier ensembles. *Sensors and Actuators B: Chemical*, 166:320–329, 2012. 28, 54, 59, 73

[226] Neil J Vickers. Mechanisms of animal navigation in odor plumes. *The Biological Bulletin*, 198(2):203–212, 2000. 12

[227] Richard G Vogt. Biochemical diversity of odor detection: Obps, odes and snmps. In *Insect pheromone biochemistry and molecular biology*, pages 391–445. Elsevier, 2003. 16

[228] Leslie B Vosshall, Allan M Wong, and Richard Axel. An olfactory sensory map in the fly brain. *Cell*, 102(2):147–159, 2000. 16

[229] Jing W Wang, Allan M Wong, Jorge Flores, Leslie B Vosshall, and Richard Axel. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell*, 112(2):271–282, 2003. 6, 17, 18

[230] Yalin Wang, Ann-Shyn Chiang, Shouzhen Xia, Toshi Kitamoto, Tim Tully, and Yi Zhong. Blockade of neurotransmission in drosophila mushroom bodies impairs odor attraction, but not repulsion. *Current Biology*, 13(21):1900–1904, 2003. 22

[231] Jan Wessnitzer, Joanna M Young, J Douglas Armstrong, and Barbara Webb. A model of non-elemental olfactory learning in drosophila. *Journal of computational neuroscience*, 32(2):197–212, 2012. 8

[232] Till Westermayer. Locust i. https://www.flickr.com/photos/98652633@N00/6200767901. CC BY-SA 2.0. 6

[233] John A White, Carson C Chow, Jason Rit, Cristina Soto-Treviño, and Nancy Kopell. Synchronization and oscillatory dynamics in heterogeneous, mutually inhibited neurons. *Journal of computational neuroscience*, 5(1):5–16, 1998. 24

[234] W Bryan Wilent and Diego Contreras. Stimulus-dependent changes in spike threshold enhance feature selectivity in rat barrel cortex neurons. *Journal of Neuroscience*, 25(11):2983–2991, 2005. 24, 25, 37

[235] Rachel I Wilson. Early olfactory processing in drosophila: mechanisms and principles. *Annual review of neuroscience*, 36:217–241, 2013. 7, 48, 49, 52, 53, 86, 94

[236] Rachel I Wilson, Glenn C Turner, and Gilles Laurent. Transformation of olfactory representations in the drosophila antennal lobe. *Science*, 303(5656):366–370, Jan 2004. 9, 25, 44, 69, 92, 99, 117

[237] Allan M Wong, Jing W Wang, and Richard Axel. Spatial representation of the glomerular map in the drosophila protocerebrum. *Cell*, 109(2):229–241, 2002. 7, 20, 22, 31

[238] Geraldine A Wright and Brian H Smith. Different thresholds for detection and discrimination of odors in the honey bee (apis mellifera). *Chem Senses*, 29(2):127–135, Feb 2004. 56

[239] Chia-Lin Wu, Meng-Fu Maxwell Shih, Jason Sih-Yu Lai, Hsun-Ti Yang, Glenn C Turner, Linyi Chen, and Ann-Shyn Chiang. Heterotypic gap junctions between two neurons in the drosophila brain are critical for memory. *Current Biology*, 21(10):848–854, 2011. 22

[240] Daniel G Wustenberg, Milena Boytcheva, Bernd Grunewald, John H Byrne, Randolf Menzel, and Douglas A Baxter. Current-and voltage-clamp recordings and computer simulations of kenyon cells in the honeybee. *Journal of neurophysiology*, 92(4):2589–2603, 2004. 8

[241] Shouzhen Xia and Tim Tully. Segregation of odor identity and intensity during odor discrimination in drosophila mushroom body. *PLoS biology*, 5(10):e264, 2007. 22

[242] Yoshiyuki Yamazaki, Michiko Nishikawa, and Makoto Mizunami. Three classes of gaba-like immunoreactive neurons in the mushroom body of the cockroach. *Brain research*, 788(1-2):80–86, 1998. 8, 22

[243] Kouji Yasuyama, Ian A Meinertzhagen, and Friedrich-Wilhelm Schürmann. Synaptic organization of the mushroom body calyx in drosophila melanogaster. *Journal of Comparative Neurology*, 445(3):211–226, 2002. 7, 30

[244] Guo-Sheng Yi, Jiang Wang, Kai-Ming Tsang, Xi-Le Wei, and Bin Deng. Input-output relation and energy efficiency in the neuron with different spike threshold dynamics. *Frontiers in computational neuroscience*, 9:62, 2015. 24, 25, 37

[245] Man Yi Yim, Ad Aertsen, and Stefan Rotter. Impact of intrinsic biophysical diversity on the activity of spiking neurons. *Physical Review E*, 87(3):032710, 2013. 24

[246] T Zars, M Fischer, R Schulz, and M Heisenberg. Localization of a short-term memory in drosophila. *Science*, 288(5466):672–675, 2000. 31

[247] Andrei Zavada, Christopher L Buckley, Dominique Martinez, Jean-Pierre Rospars, and Thomas Nowotny. Competition-based model of pheromone component ratio detection in the moth. *PLoS One*, 6(2):e16308, 2011. 22, 23, 24, 25, 44, 61, 91, 98